PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/00 — A2

(11) International Publication Number: **WO 99/09151**

(43) International Publication Date: 25 February 1999 (25.02.99)

(21) International Application Number: PCT/US98/14841

(22) International Filing Date: 17 July 1998 (17.07.98)

(30) Priority Data:
08/910,386 13 August 1997 (13.08.97) US

(71) Applicant: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA [US/US]; Office of Technology Transfer, 5th floor, 1111 Franklin Street, Oakland, CA 94607–5200 (US).

(72) Inventors: RONALD, Pamela, C.; 26951 Road 96, Davis, CA 95616 (US). WANG, Guo–Liang; 927 J Street #50, Davis, CA 95616 (US). SONG, Wen–Yuang; 22A Salano Park, Davis, CA 95616 (US). SZABO, Veronique; 1880 Cowell Boulevard, Davis, CA 95616 (US). HULBERT, Scot, H.; 319 Denison Street, Manhattan, KS 66516 (US). RICHTER, Todd; 1211 Devon Court, Kokomo, IN 94609 (US).

(74) Agents: BASTIAN, Kevin, L. et al.; Townsend and Townsend and Crew LLP, 8th floor, Two Embarcadero Center, San Francisco, CA 94111–3834 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: PROCEDURES AND MATERIALS FOR CONFERRING DISEASE RESISTANCE IN PLANTS

(57) Abstract

The present invention provides nucleic acids encoding polypeptides which confer resistance to *Xanthomonas* spp and other pathogens. The nucleic acids can be used to produce transgenic plants resistant to the pathogen.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

PROCEDURES AND MATERIALS FOR CONFERRING DISEASE RESISTANCE IN PLANTS

This application is related to U.S. Patent Application No. 08/587,680, filed January 17, 1996, which is a continuation in part of copending U.S. patent application No. 08/567,375, filed December 4, 1995, which is a continuation in part of U.S. provisional patent application No., 60/004,645. The '680 application is also a continuation in part of copending U.S. patent application No. 08/475,891, filed June 7, 1995, which is a continuation in part of copending U.S. patent application No. 08/373,374, filed January 17, 1995. These applications are incorporated herein by reference.

Field Of The Invention

The present invention relates generally to plant molecular biology. In particular, it relates to nucleic acids and methods for conferring disease resistance in plants.

Statement as to Rights to Inventions Made Under Federally Sponsored Research and Development

This invention was made with Government support under Grant No. GM47907, awarded by the National Institutes of Health and Grant No. 9300834, awarded by the United States Department of Agriculture. The Government has certain rights in this invention.

BACKGROUND OF THE INVENTION

Loci conferring disease resistance have been identified in many plant species. Genetic analysis of many plant-pathogen interactions has demonstrated that plants contain loci that confer resistance against specific races of a pathogen containing a complementary avirulence gene. Molecular characterization of these genes should provide means for conferring disease resistance to a wide variety of crop plants.

Those plant resistance genes that have been characterized at the molecular level fall into four classes. One gene, *Hm1* in corn, encodes a reductase and is effective against the fungal pathogen *Cochliobolus carbonum* (Johal *et al. Science* 258:985-987 (1992)). In tomato, the *Pto* gene confers resistance against *Pseudomonas syringae* that express the *avrPto* avirulence gene (Martin *et al. Science* 262:1432 (1993)). The predicted *Pto* gene encodes a serine threonine protein kinase. The tomato *Cf-9* gene confers resistance to races of the fungus *Cladosporium fulvum* that carry the avirulence gene *Avr9* (Jones *et al. Science* 266:789- 793 (1994). The tomato *Cf-9* gene encodes a putatitive extracellular LRR protein. Finally, the *RPS2* gene of *Arabidopsis thaliana* confers resistance to *P. syringae* that express the *avrRpt2* avirulence gene (Bent *et al. Science* 265:1856-1860 (1994)). *RPs2* encodes a protein with an LRR motif and a P-loop motif.

Bacterial blight disease caused by *Xanthomonas* spp. infects virtually all crop plants and leads to extensive crop losses worldwide. Bacterial blight disease of rice (*Oryza sativa*), caused by *Xanthomonas oryzae* pv. *oryzae* (*Xoo*), is an important disease of this crop. Races of *Xoo* that induce resistant or susceptible reactions on rice cultivars with distinct resistance (*Xa*) genes have been identified. One source of resistance (*Xa21*) had been identified in the wild species *Oryza longistaminata* (Khush *et al. in Proceedings of the International Workshop on Bacterial Blight of Rice.* (International Rice Research Institute, 1989) and Ikeda *et al. Jpn J. Breed* 40 (Suppl.1):280-281 (1990)). *Xa21* is a dominant resistance locus that confers resistance to all known isolates of *Xoo* and is the only characterized *Xa* gene that carries resistance to *Xoo* race 6. Genetic and physical analysis of the *Xa21* locus has identified a number of tightly linked markers on chromosome 11 (Ronald *et al. Mol. Gen. Genet.* 236:113-120 (1992)). The molecular mechanisms by which the *Xa21* locus confers resistance to this pathogen were not identified, however.

Considerable effort has been directed toward cloning plant genes conferring resistance to a variety of bacterial, fungal and viral diseases. Only one pest resistance gene has been cloned in monocots. Since monocot crops feed most humans and animals in the world, the identification of disease resistance genes in these plants is particularly important. The present invention addresses these and other needs.

SUMMARY OF THE INVENTION

The present invention provides isolated nucleic acid constructs comprising an *RRK* polynucleotide sequence. The sequences can be rice sequences which hybridize to SEQ ID NOs: 1, 4, 6, 8, 10, or 11 under stringent conditions. Also claimed are sequences from cassava which hybdridize to SEQ ID NO: 13), maize sequences which hybridize to SEQ ID NOs: 15, 16), and tomato (*e.g.*, SEQ ID NOs:17, 19, or 21). Exemplary *RRK* polynucleotide sequences are *Xa21* sequences which encode an *Xa21* polypeptide as shown below. The RRK polynucleotides encode a protein having a leucine rich repeat motif and/or a cytoplasmic protein kinase domain. The nucleic acid constructs of the invention may further comprise a promoter operably linked to the *RRK* polynucleotide sequence. The promoter may be a tissue-specific promoter or a constitutive promoter.

The invention also provides nucleic acid constructs comprising a promoter sequence from an *RRK* gene linked to a heterologous polynucleotide sequence. Exemplary heterologous polynucleotide sequences include structural genes which confer pathogen resistance on plants.

The invention further provides transgenic plants comprising a recombinant expression cassette comprising a promoter from an *RRK* gene operably linked to a polynucleotide sequence as well as transgenic plants comprising a recombinant expression cassette comprising a plant promoter operably linked to an *RRK* polynucleotide sequence. Although any plant can be used in the invention, rice and tomato plants may be conveniently used.

The invention further provides methods of enhancing resistance to *Xanthomonas* and other pathogens in a plant. The methods comprise introducing into the plant a recombinant expression cassette comprising a plant promoter operably linked to an RRK polynucleotide sequence. The methods may be conveniently carried out with rice or tomato plants.

Definitions

The term "plant" includes whole plants, plant organs (e.g., leaves, stems, roots, etc.), seeds and plant cells and progeny of same. The class of plants which can be used in the methods of the invention is generally as broad as the class of higher plants

amenable to transformation techniques, including both monocotyledonous and dicotyledonous plants.

A "heterologous sequence" is one that originates from a foreign species, or, if from the same species, is substantially modified from its original form. For example, a promoter operably linked to a heterologous structural gene is from a species different from that from which the structural gene was derived, or, if from the same species, one or both are substantially modified from their original form.

An "*RRK* gene" is member of a new class of disease resistance genes which encode RRK polypeptides which typically comprise an extracellular LRR domain, a transmembrane domain, and a cytoplasmic protein kinase domain (as shown in *e.g.*, *Pto* and *Fen* (Martin *et al. Plant Cell* 6:1543-1552 (1994)). As used herein, an LRR domain is a region of a repeated unit of about 24 residues as described in USSN 08/587,680, and found in *Cf-9*). Using the sequences disclosed here and standard nucleic acid hybridization and/or amplification techniques, one of skill can identify members of this class of genes. For instance, a nucleic acid probe from an *Xa21* gene detected polymorphisms that segregated with the blast (*Pyricularia oryzae*) resistance gene (Pi7) in 58 recombinant inbred lines of rice. The same probe also detected polymorphism in nearly isogenic lines carrying *xa5* and *Xa10* resistance genes.

In some preferred embodiments, members of this class of disease resistance genes can be identified by their ability to be amplified by degenerate PCR primers which correspond to the LRR and kinase domains. For instance, primers have been used to isolate homologous genes in tomato, maize and cassava. The maize gene disclosed here has been genetically mapped to a region associated with resistance to *Helminthosporium turcicum*. Exemplary primers for this purpose are tcaagcaacaatttgtcaggnca (a/g) at (a/c/t) cc (for the LRR domain sequence GQIP) and taacagcacattgcttgatttnan (g/a) tcncg (g/a) tg (the kinase domain sequence HCDIK). These or equivalent primers are then used to amplify the appropriate nucleic acid using the PCR conditions described below.

An "*Xa21* polynucleotide sequence" is a subsequence or full length polynucleotide sequence of an *Xa21* gene, such as the rice *Xa21* gene, which, when present in a transgenic plant confers resistance to *Xanthomonas* spp. (*e.g.*, *X. oryzae*) on the plant. Exemplary polynucleotides of the invention include the coding region of the sequences provided below. An *Xa21* polynucleotide is typically at least about 3100

nucleotides to about 6500 nucleotides in length, usually from about 4000 to about 4500 nucleotides.

An "*Xa21* polypeptide" is a gene product of an *Xa21* polynucleotide sequence, which has the activity of *Xa21*, *i.e.*, the ability to confer resistance to *Xanthomonas spp*. *Xa21* polypeptides, like other *RRK* polypeptides, are characterized by the presence of an extracellular domain comprising a region of leucine rich repeats (LRR) and/or a cytoplasmic protein kinase domain. Exemplary *Xa21* polypeptides of the invention include those described below.

In the expression of transgenes one of skill will recognize that the inserted polynucleotide sequence need not be identical and may be "substantially identical" to a sequence of the gene from which it was derived. As explained below, these variants are specifically covered by this term.

In the case where the inserted polynucleotide sequence is transcribed and translated to produce a functional *RRK* polypeptide, one of skill will recognize that because of codon degeneracy, a number of polynucleotide sequences will encode the same polypeptide. These variants are specifically covered by the term "*RRK* polynucleotide sequence". In addition, the term specifically includes those full length sequences substantially identical (determined as described below) with an *RRK* gene sequence and that encode proteins that retain the function of the *RRK* protein. Thus, in the case of rice *RRK* genes disclosed here, the above term includes variant polynucleotide sequences which have substantial identity with the sequences disclosed here and which encode proteins capable of conferring resistance to *Xanthomonas* or other plant diseases and pests on a transgenic plant comprising the sequence.

Two polynucleotides or polypeptides are said to be "identical" if the sequence of nucleotides or amino acid residues, respectively, in the two sequences is the same when aligned for maximum correspondence as described below. The term "complementary to" is used herein to mean that the complementary sequence is identical to all or a portion of a reference polynucleotide sequence.

Sequence comparisons between two (or more) polynucleotides or polypeptides are typically performed by comparing sequences of the two sequences over a segment or "comparison window" to identify and compare local regions of sequence similarity. The segment used for purposes of comparison may be at least about 20

contiguous positions, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned.

Optimal alignment of sequences for comparison may be conducted by the local homology algorithm of Smith and Waterman *Adv. Appl. Math.* 2: 482 (1981), by the homology alignment algorithm of Needleman and Wunsch *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson and Lipman *Proc. Natl. Acad. Sci. (U.S.A.)* 85: 2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), or by inspection.

"Percentage of sequence identity" is determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity.

The term "substantial identity" of polynucleotide sequences means that a polynucleotide comprises a sequence that has at least 60% sequence identity, preferably at least 80%, more preferably at least 90% and most preferably at least 95%, compared to a reference sequence using the programs described above (preferably BESTFIT) using standard parameters. One of skill will recognize that these values can be appropriately adjusted to determine corresponding identity of proteins encoded by two nucleotide sequences by taking into account codon degeneracy, amino acid similarity, reading frame positioning and the like. Substantial identity of amino acid sequences for these purposes normally means sequence identity of at least 40%, preferably at least 60%, more preferably at least 90%, and most preferably at least 95%. Polypeptides which are "substantially similar" share sequences as noted above except that residue positions which are not identical may differ by conservative amino acid changes. Conservative amino acid substitutions refer to the interchangeability of residues having similar side chains. For

example, a group of amino acids having aliphatic side chains is glycine, alanine, valine, leucine, and isoleucine; a group of amino acids having aliphatic-hydroxyl side chains is serine and threonine; a group of amino acids having amide-containing side chains is asparagine and glutamine; a group of amino acids having aromatic side chains is phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains is lysine, arginine, and histidine; and a group of amino acids having sulfur-containing side chains is cysteine and methionine. Preferred conservative amino acids substitution groups are: valine-leucine-isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-valine, and asparagine-glutamine.

Another indication that nucleotide sequences are substantially identical is if two molecules hybridize to each other under appropriate conditions. Appropriate conditions can be high or low stringency and will be different in different circumstances. Generally, stringent conditions are selected to be about 5°C to about 20°C lower than the thermal melting point (Tm) for the specific sequence at a defined ionic strength and pH. The Tm is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Typically, stringent wash conditions are those in which the salt concentration is about 0.02 molar at pH 7 and the temperature is at least about 60°C. However, nucleic acids which do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This may occur, *e.g.*, when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. For Southern hybridizations, high stringency wash conditions will include at least one wash in 0.1X SSC at 65°C.

Nucleic acids of the invention can be identified from a cDNA or genomic library prepared according to standard procedures and the nucleic acids disclosed here (typically at least 100 nucleotides to about full length) used as a probe. Low stringency hybridization conditions will typically include at least one wash using 2X SSC at 65°C. The washes are preferrably followed by a subsequent wash using 1X SSC at 65°C.

As used herein, a homolog of a particular *RRK* gene (*e.g.*, the rice *Xa21* genes disclosed here) is a second gene (either in the same species or in a different species) which encodes a protein having an amino acid sequence having at least 25% identity or 45% similiarity to (determined as described above) to a polypeptide sequence in the first

gene product. It is believed that, in general, homologs share a common evolutionary past.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the genome organization of the seven *Xa21* family members and location of 14 transposon-like elements. Cosmid and BAC clones carrying the family members are designated. Wide bars represent predicted coding regions, fine bars represent noncoding regions, introns are indicated by angled lines, and the non-sequenced regions are shown by straight lines. A gap in the sequence of BAC9 is indicated by "//". Letters refer to names of *Xa21* gene family members and arrows indicate direction of ORFs. The 14 transposon-like elements are numbered and represented by closed triangles.

Figure 2A shows the HC region of the sequenced *Xa21* gene family members. Wide bars represent predicted coding regions, and fine bars represent non-coding regions. Start and stop codons are indicated. The 5' flanking regions and downstream regions are grouped into four and two groups, respectively, and are shown in different colors based on sequence identity. The percentage of DNA sequence identity between promoter regions and between classes is shown to the left and right, respectively. The HC region is indicated by a black bar.

Figure 2B is a schematic diagram showing a comparison of the predicted amino acid sequences of XA21 and A1. Domains are numbered as follows: I, Presumed signal peptide; II, presumed N terminus; III, LRR; VI, charged; V, presumed transmembrane; VI charged; VII juxtamembrane; VIII, serine/threonine kinase; IX, carboxy tail. The numbers below each domain indicate amino acid identity between XA21 and A1.

Figure 3A shows family member D and insertion position of *Retrofit*. *Retrofit* carries long terminal repeats (LTRs) (small arrows) and a single, large ORF, encoding a protein with the following domains: gag, protease (PR), integrase (IN), reverse transcriptase (RT), and RNase H (RH). The large arrow indicates direction of the ORF.

Figure 3B shows family member E and insertion position of *Truncator*. Arrows mark the orientation of the inverted repeats. The deduced amino acid sequences of the tomato resistance genes Cf9 and Pto are shown below. In both Figures 3A and 3B, the insertion elements are designated by a hatched bar. The presumed deduced amino acid sequences of members D and E are shown by shaded rectangles. Domains representations

are as described in the legend to Figure 2.

Figure 4 shows intergenic recombination break point in the *Xa21* family members. Boxes represent the ORFs of the designated family members, while narrow boxes represent flanking regions. Same colors indicate a high level of sequence homology. The nucleotides of the presumed recombination break points are indicated in large and bold type. Sequences surrounding the recombination break point are also shown.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

This invention relates to plant *RRK* genes, such as the *Xa21* genes of rice. Nucleic acid sequences from *RRK* genes, in particular *Xa21* genes, can be used to confer resistance to *Xanthomonas* and other pathogens in plants. The invention has use in conferring resistance in all higher plants susceptible to pathogen infection. The invention thus has use over a broad range of types of plants, including species from the genera *Juglans, Fragaria, Lotus, Medicago, Onobrychis, Trifolium, Trigonella, Vigna, Citrus, Linum, Geranium, Manihot, Daucus, Arabidopsis, Brassica, Raphanus, Sinapis, Atropa, Capsicum, Datura, Hyoscyamus, Lycopersicon, Nicotiana, Solanum, Petunia, Digitalis, Majorana, Ciahorium, Helianthus, Lactuca, Bromus, Asparagus, Antirrhinum, Heterocallis, Nemesis, Pelargonium, Panieum, Pennisetum, Ranunculus, Senecio, Salpiglossis, Cucumis, Browaalia, Glycine, Pisum, Phaseolus, Lolium, Zea, Avena, Hordeum, Secale, Triticum,* and, *Sorghum.*

The Example section below, which describes the isolation and characterization of *RRK* genes in rice, casava, maize and tomato. The methods used to isolate these genes are exemplary of a general approach for isolating *Xa21* genes and other *RRK* genes. The isolated genes can then be used to construct recombinant vectors for transferring *RRK* gene expression to transgenic plants.

Generally, the nomenclature and the laboratory procedures in recombinant DNA technology described below are those well known and commonly employed in the art. Standard techniques are used for cloning, DNA and RNA isolation, amplification and purification. Generally enzymatic reactions involving DNA ligase, DNA polymerase, restriction endonucleases and the like are performed according to the manufacturer's specifications. These techniques and various other techniques are generally performed

according to Sambrook *et al.*, *Molecular Cloning - A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, (1989).

The isolation of *Xa21* and related *RRK* genes may be accomplished by a number of techniques. For instance, oligonucleotide probes based on the sequences disclosed here can be used to identify the desired gene in a cDNA or genomic DNA library. To construct genomic libraries, large segments of genomic DNA are generated by random fragmentation, e.g. using restriction endonucleases, and are ligated with vector DNA to form concatemers that can be packaged into the appropriate vector. To prepare a cDNA library, mRNA is isolated from the desired organ, such as leaf and a cDNA library which contains the *RRK* gene transcript is prepared from the mRNA. Alternatively, cDNA may be prepared from mRNA extracted from other tissues in which RRK genes or homologs are expressed.

The cDNA or genomic library can then be screened using a probe (typically a degenerate probe) based upon the sequence of a cloned *RRK* gene such as rice *Xa21* genes disclosed here. Probes may be used to hybridize with genomic DNA or cDNA sequences to isolate homologous genes in the same or different plant species.

Alternatively, the nucleic acids of interest can be amplified from nucleic acid samples using amplification techniques. For instance, polymerase chain reaction (PCR) technology to amplify the sequences of the *RRK* and related genes directly from genomic DNA, from cDNA, from genomic libraries or cDNA libraries. PCR and other *in vitro* amplification methods may also be useful, for example, to clone nucleic acid sequences that code for proteins to be expressed, to make nucleic acids to use as probes for detecting the presence of the desired mRNA in samples, for nucleic acid sequencing, or for other purposes.

Appropriate primers and probes for identifying *RRK* sequences from plant tissues are generated from comparisons of the sequences provided herein. For a general overview of PCR see *PCR Protocols: A Guide to Methods and Applications*. (Innis, M, Gelfand, D., Sninsky, J. and White, T., eds.), *Academic Press*, San Diego (1990), incorporated herein by reference.

Polynucleotides may also be synthesized by well-known techniques as described in the technical literature. See, e.g., Carruthers *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* 47:411-418 (1982), and Adams *et al.*, *J. Am. Chem. Soc.* 105:661

(1983). Double stranded DNA fragments may then be obtained either by synthesizing the complementary strand and annealing the strands together under appropriate conditions, or by adding the complementary strand using DNA polymerase with an appropriate primer sequence.

Isolated sequences prepared as described herein can then be used to provide *RRK* gene expression and therefore *Xanthomonas* resistance in desired plants. One of skill will recognize that the nucleic acid encoding a functional *RRK* protein need not have a sequence identical to the exemplified gene disclosed here. In addition, the polypeptides encoded by the *RRK* genes, like other proteins, have different domains which perform different functions. Thus, the *RRK* gene sequences need not be full length, so long as the desired functional domain of the protein is expressed. As explained in detail below, the proteins of the invention comprise an extracellular leucine rich repeat domain, as well as an intracellular kinase domain. Modified protein chains can also be readily designed utilizing various recombinant DNA techniques well known to those skilled in the art. For example, the chains can vary from the naturally occurring sequence at the primary structure level by amino acid substitutions, additions, deletions, and the like. Modification can also include swapping domains from the proteins of the invention with related domains from other pest resistance genes. For example, the extra cellular domain (including the leucine rich repeat region) of the proteins of the invention can be replaced by that of the tomato *Cf-9* gene and thus provide resistance to fungal pathogens of rice. These modifications can be used in a number of combinations to produce the final modified protein chain.

To use isolated *RRK* sequences in the above techniques, recombinant DNA vectors suitable for transformation of plant cells are prepared. Techniques for transforming a wide variety of higher plant species are well known and described in the technical and scientific literature. See, for example, Weising *et al. Ann. Rev. Genet.* 22:421-477 (1988).

A DNA sequence coding for the desired *RRK* polypeptide, for example a cDNA or a genomic sequence encoding a full length protein, will be used to construct a recombinant expression cassette which can be introduced into the desired plant. An expression cassette will typically comprise the *RRK* polynucleotide operably linked to transcriptional and translational initiation regulatory sequences which will direct the

transcription of the sequence from the *RRK* gene in the intended tissues of the transformed plant.

For example, a plant promoter fragment may be employed which will direct expression of the *RRK* in all tissues of a regenerated plant. Such promoters are referred to herein as "constitutive" promoters and are active under most environmental conditions and states of development or cell differentiation. Examples of constitutive promoters include the cauliflower mosaic virus (CaMV) 35S transcription initiation region, the 1'- or 2'-promoter derived from T-DNA of *Agrobacterium tumafaciens*, and other transcription initiation regions from various plant genes known to those of skill.

Alternatively, the plant promoter may direct expression of the *RRK* gene in a specific tissue or may be otherwise under more precise environmental or developmental control. Such promoters are referred to here as "inducible" promoters. Examples of environmental conditions that may effect transcription by inducible promoters include pathogen attack, anaerobic conditions, or the presence of light.

Examples of promoters under developmental control include promoters that initiate transcription only in certain tissues, such as leaves, roots, fruit, seeds, or flowers. The operation of a promoter may also vary depending on its location in the genome. Thus, an inducible promoter may become fully or partially constitutive in certain locations.

The endogenous promoters from the *RRK* genes of the invention can be used to direct expression of the genes. These promoters can also be used to direct expression of heterologous structural genes. Thus, the promoters can be used in recombinant expression cassettes to drive expression of genes conferring resistance to any number of pathogens, including fungi, bacteria, and the like.

To identify the promoters, the 5' portions of the clones described here are analyzed for sequences characteristic of promoter sequences. For instance, promoter sequence elements include the TATA box consensus sequence (TATAAT), which is usually 20 to 30 base pairs upstream of the transcription start site. In plants, further upstream from the TATA box, at positions -80 to -100, there is typically a promoter element with a series of adenines surrounding the trinucleotide G (or T) N G. J. Messing et al., in *Genetic Engineering in Plants*, pp. 221-227 (Kosage, Meredith and Hollaender, eds. 1983).

If proper polypeptide expression is desired, a polyadenylation region at the 3'-end of the *RRK* coding region should be included. The polyadenylation region can be derived from the natural gene, from a variety of other plant genes, or from T-DNA.

The vector comprising the sequences from an *RRK* gene will typically comprise a marker gene which confers a selectable phenotype on plant cells. For example, the marker may encode biocide resistance, particularly antibiotic resistance, such as resistance to kanamycin, G418, bleomycin, hygromycin, or herbicide resistance, such as resistance to chlorosluforon or Basta.

Such DNA constructs may be introduced into the genome of the desired plant host by a variety of conventional techniques. For example, the DNA construct may be introduced directly into the genomic DNA of the plant cell using techniques such as electroporation, PEG poration, particle bombardment and microinjection of plant cell protoplasts or embryogenic callus, or the DNA constructs can be introduced directly to plant tissue using ballistic methods, such as DNA particle bombardment. Alternatively, the DNA constructs may be combined with suitable T-DNA flanking regions and introduced into a conventional *Agrobacterium tumefaciens* host vector. The virulence functions of the *Agrobacterium tumefaciens* host will direct the insertion of the construct and adjacent marker into the plant cell DNA when the cell is infected by the bacteria.

Transformation techniques are known in the art and well described in the scientific and patent literature. The introduction of DNA constructs using polyethylene glycol precipitation is described in Paszkowski *et al. Embo J.* 3:2717-2722 (1984). Electroporation techniques are described in Fromm *et al. Proc. Natl. Acad. Sci. USA* 82:5824 (1985). Ballistic transformation techniques are described in Klein *et al. Nature* 327:70-73 (1987). Using a number of approaches, cereal species such as rye (de la Pena *et al.*, *Nature* 325:274-276 (1987)), corn (Rhodes *et al.*, *Science* 240:204-207 (1988)), and rice (Shimamoto *et al.*, *Nature* 338:274-276 (1989) by electroporation; Li *et al. Plant Cell Rep.* 12:250-255 (1993) by ballistic techniques) can be transformed.

Agrobacterium tumefaciens-meditated transformation techniques are well described in the scientific literature. See, for example Horsch *et al. Science* 233:496-498 (1984), and Fraley *et al. Proc. Natl. Acad. Sci. USA* 80:4803 (1983). Although *Agrobacterium* is useful primarily in dicots, certain monocots can be transformed by *Agrobacterium*. For instance, *Agrobacterium* transformation of rice is described by Hiei *et*

al, *Plant J.* 6:271-282 (1994).

Transformed plant cells which are derived by any of the above transformation techniques can be cultured to regenerate a whole plant which possesses the transformed genotype and thus the desired *RRK*-controlled phenotype. Such regeneration techniques rely on manipulation of certain phytohormones in a tissue culture growth medium, typically relying on a biocide and/or herbicide marker which has been introduced together with the *RRK* nucleotide sequences. Plant regeneration from cultured protoplasts is described in Evans *et al.*, *Protoplasts Isolation and Culture*, *Handbook of Plant Cell Culture*, pp. 124-176, MacMillilan Publishing Company, New York, 1983; and Binding, *Regeneration of Plants*, *Plant Protoplasts*, pp. 21-73, CRC Press, Boca Raton, 1985. Regeneration can also be obtained from plant callus, explants, organs, or parts thereof. Such regeneration techniques are described generally in Klee *et al. Ann. Rev. of Plant Phys.* 38:467-486 (1987).

The methods of the present invention are particularly useful for incorporating the *RRK* polynucleotides into transformed plants in ways and under circumstances which are not found naturally. In particular, the *RRK* polypeptides may be expressed at times or in quantities which are not characteristic of natural plants.

One of skill will recognize that after the expression cassette is stably incorporated in transgenic plants and confirmed to be operable, it can be introduced into other plants by sexual crossing. Any of a number of standard breeding techniques can be used, depending upon the species to be crossed.

The effect of the modification of *RRK* gene expression can be measured by detection of increases or decreases in mRNA levels using, for instance, Northern blots. In addition, the phenotypic effects of gene expression can be detected by measuring lesion length as in plants. Suitable assays for determining resistance are described in USSN 08/587,680.

The following Examples are offered by way of illustration, not limitation.

Example 1

As noted above, *Xa21* genes make up a multigene family. Pulsed field gel electrophoresis and genetic analysis have demonstrated that most of the members of the *Xa21* gene family are located in a 230 kb genomic region on chromosome 11 linked to at

least 8 major resistance genes and 1 QTL for resistance (Song, *et al.*, *Science* 270:1804 (1995); Ronald, *et al.*, *Mol. Gen. Genet.* 236:113 (1992).

This example describes six *Xa21* gene family members from the resistant rice line IRBB21, which members are designated A1, A2, C, D, E, and F. Cloning was as described in USSN 08/587,680; Song, *et al.*, *supra* and Wang, *et al.*, *Plant J.* 7, 525 (1995). DNA sequences were determined by using the Sequitherm Long Read Cycle Sequencing Kit (Epicentre Technologies) in combination with the LI-COR Model 4000L Automated Sequencer (LI-COR Inc). To fill in gaps, a primer walking strategy was performed using synthesized primers (Operon) and the Applied Biosystems 373 DNA sequencer. Genebank accession numbers are as follows: A1: U72725 (SEQ ID NO: 4); A2: U72727 (SEQ ID NO: 10); C: U72723 (SEQ ID NO: 6); D: U72726 (SEQ ID NO: 1); E: U72724 (SEQ ID NO: 8); F: U72728 (SEQ ID NO: 12); 3' flanking region of F: U72729 (SEQ ID NO: 12). The Wisconsin sequence analysis programs GAP and Pileup were used to calculate the percent identity and to carry out multiple alignments of DNA and protein sequences, respectively.

Sequence data and restriction enzyme analysis of cosmid and bacterial artificial chromosome clones indicated that the seven members are contained on 4 clones (Fig. 1). The first clone, carrying *Xa21* (described in USSN 08/587,680 and Song *et al.*, *supra*. The Genbank accession number for *Xa21* genmomic and cDNA sequences is U37133) and member C, spans a 40 kb region; the second clone includes member D, A1, and A2 and occupies a 150 kb region; clones of 40 kb and 130 kb contain members E and F, respectively. Genetic and molecular data suggests member E is inherited from the susceptible parent IR24 (P.C. Ronald, *et al.*, *Mol. Gen. Genet.* **236**, 113 (1992)).

The entire coding region, the intron, and 3' flanking region of the seven family members can be grouped into two classes. One class (designated the *Xa21* class) contains *Xa21*, as well as members D and F (SEQ ID NOs: 1 and 12). The second class (designated the A2 class) contains members A1 (SEQ ID NO:4), A2 (SEQ ID NO:10), C (SEQ ID NO:6), and E (SEQ ID NO:8). Within each class, family members share striking nucleotide sequence identity (98.0% average identity for the members of the *Xa21* class; 95.2% average identity for the members of the A2 class); compared to low levels of DNA sequence identity between members of the two classes (eg. 63.5% identity between *Xa21* and A2) (Fig. 2A). Only the *Xa21* and A1 open reading frames (ORFs) encode

receptor kinase-like proteins. The sequence of other family members contain alterations causing a premature truncation of the predicted receptor kinase-like ORF (small deletions in F and C; base pair mutations in A2; or transposon insertions in D and E). At the amino acid level, Al and XA21 share 68.6% identity overall. As shown in Figure 2B, Domains I and II, carrying the presumed signal peptide and amino terminus of the protein, are 100% identical whereas the LRR domain (domain III) of XA21 and A1 share a low level of identity (59.5%) and differ in the number of LRRs (23 vs 22 respectively). In the presumed intracellular portion, the catalytic domains (domain VIII) of XA21 and Al are highly conserved (82% identity), whereas the non-catalytic regions are divergent (64% identity for domain VII (juxtamembrane) and 38.5% identity for domain IX (carboxyl terminus)). The differences observed between members of the two classes suggest that they may differ in function. Indeed, we have found transgenic plants containing the A1 sequence are susceptible to all *Xoo* isolates tested.

A remarkable feature of the *Xa21* family members is the presence of fourteen transposable element-like sequences (M.A. Grandbastien, *et al.*, *Nature* **337:** 376 (1989); S.E.; White, *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **91:** 11792 (1994)). The position of these elements is shown in Fig 1. Twelve elements insert into noncoding regions; whereas two elements, named *Retrofit* and *Truncator*, integrate into the coding regions of members D and E, respectively, resulting in disruption of the ORFs of these two members (Fig. 1, number 9 and 13). *Retrofit* (SEQ ID NO:3) belongs to the *Drosophila* copia class of retrotransposons and carries a large ORF showing greatest similarity to the ORF of maize *Hopscotch* (68.6% similarity; 54.6% identity) and tobacco *Tnt1* (51.4% similarity; 31.9% identity) (M.A. Grandbastien, *et al.*, *Nature* **337:** 376 (1989); S.E.; White, *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **91:** 11792 (1994)). The insertion site of this element is located between the 23rd (V) and 24th (P) amino acids of the 22nd LRR creating a truncated molecule, lacking the transmembrane and kinase domains (Fig. 3A). Insertion of *Retrofit* into a presumed coding region contrasts with the observation in yeast and maize that integration of retrotransposons is biased towards noncoding regions (D.F. Voytas, *Science* **274:** 737 (1996); P. SanMiguel, *et al.*, *Science* **274:** 765 (1996)). The fact that the truncated D confers partial resistance to *Xoo* suggests that transposition events at the *Xa21* locus can alter expression of resistance.

Truncator, 2913 bp, represents a novel transposon-like sequence carrying 9

bp terminal inverted repeats (TIRs). The sequence shows no significant homology to any sequence in the database and contains no obvious ORFs. Interestingly, insertion of this element into the amino terminus of the kinase domain of member E would presumably result in premature truncation of the receptor kinase resulting in a receptor-like molecule structurally similar to the tomato fungal resistance gene products Cf9 and Cf2 (Fig. 3B) (D.A. Jones, *et al.*, *Science*, **266**: 789 (1994); M.S. Dixon, *et al.*, *Cell* 84:451 (1996)).

In addition to the transposition events presented above, recombination between different family members was also found to play an important role in the evolution of the *Xa21* locus. A 269 bp highly conserved (HC) region, located immediately downstream of the start codon of all seven family members marks the site of intragenic recombination events (Fig. 2A). The HC region, has a high G/C content (61.8% for *Xa21*) hallmarked by the typical G/C rich restriction enzyme recognition site Not I. At the amino acid level, the HC region spans domain I and domain H of XA21 and shares nearly 100% identity among seven family members.

The HC region delimits four classes of DNA sequences (~1.3 kb) upstream of the HC region. The 5' flanking region of family member F is divergent from that of other family members (less than 40% identity). The precise breakpoint (from sequence similarity to divergence) between *Xa21* and F is located within the HC region, 120 bp downstream from the start codon. This sudden change of sequence identity is unlikely due to random events such as transposon insertion or deletion because such events would presumably lead to an altered coding region. This is not the case; the deduced amino acid sequence of F maintains the receptor kinase like ORF. These results suggest that a recombination event occurred in the HC region resulting in the formation of a chimeric sequence containing the 5' flanking region of F and a downstream region (including coding region, intron, and 3' flanking region) of the *Xa21* class.

In further support of the idea that the HC region mediates intragenic recombination, we also observed apparent recombination breakpoints near or within the HC region for gene family members E, A1, and C. For E, the 5' flanking region is divergent from all other members whereas the 3' downstream regions belong to the A2 class. The sudden change of DNA identity can be explained by a recombination event between a progenitor A2-type gene and an unknown family member. The likely recombination breakpoint in E is located 105 bp upstream of the HC region since

sequences upstream of this site are quite different, compared with a high level of DNA sequence identity downstream of this site.

The nearly identical DNA sequences of C and A1 provide the most striking example of an HC mediated recombination event. For example, the 5' flanking region of C shows nearly perfect identity (99.2%) to that of *Xa21*, whereas the downstream region of C belongs to the A2 class. The high level of identity between the 5' flanking sequences of *Xa21* and C extends 3.8 kb upstream. This upstream region includes the functional promoter for the *Xa21* gene (W.-Y. Song, *et al.*, *Science* **270**:1804 (1995)). These results strongly suggest that C was created by a recombination event in the HC region between progenitors of the *Xa21* and A2 classes. The likely recombination breakpoint in member C is delimited by two characteristic deletions: one is located at position -37 and is only present in *Xa21* class members *(Xa21*, D, C, and A1); another deletion is located at position 255 and occurs in all A2 class members.

From these results it is clear that we have identified a highly conserved, G/C rich region in the gene family and that this region appears to be involved in high frequency recombination between family members. Not only is the HC region present in *O. longistaminata,* but is also present in *Xa21* family members of the cultivated rice species *O. sativa* (The clone RG103, spanning the HC region of an *Xa21* gene family member was isolated from *0. sativa* cultivar IR36 (3, S. Mcouch, *et al.*, *Theoret. Appl. Genet.* **76**:815 (1988)). Genebank accession number of RG103 is U82168. The mechanism for HC region-mediated recombination is unknown; however, two models can be envisioned. First, this region may mediate programmed recombination similar to that observed in African trypanosomes (R.H.A. Plasterk, *Trends Genet* **8**, 403 (1992)). In trypanosomes, antigenic variation is controlled by a variant surface glycoprotein (VSG), which is encoded by a member of a multigene family containing more than 1000 members. Recombination at stretches of highly conserved nucleotides between silent and expressed members of the VSG gene family leads to expression of new antigens. Alternatively, HC mediated recombination may be an example of an ectopic recombination event where the HC region serves as a recombination initiation site (T.D. Petes, *et al.*, *Annu. Rev. Genet.* **22**:147 (1988); A. Nicolas, *et al.*, *Nature* **338**: 35 (1989)). Frequent recombination in this region would maintain the conservation of the HC region but allow flanking sequences to diverge. Over time, mismatch repair would lead to homogenization of the HC region and

result in an overall increased G/C content as has been observed in yeast (Brown T., *et al.*, *Cell* 54, 705 (1988)).

Evidence for recombination in intergenic regions of the *Xa21* family members was also observed. First, sequences in the 5' flanking region of members C and *Xa21* are identical for 3.8 kb and then abruptly diverge. Interestingly, the same site of divergence is observed in the 3' flanking regions of *Xa21* and member F (Fig. 4). The presence of a conserved site of divergence suggests not only that this is a recombination breakpoint but that the *Xa21*/C cluster and member F are generated from the same progenitor. Second, the sequence of a 14742 bp region spanning the *Xa21*/C cluster shows 97.7% identity to the corresponding sequence (14871 bp) of the D/A1/A2 cluster (Fig. 1), suggesting these regions evolved through sequence duplication. This duplication process can be explained by a presumed unequal cross-over event in the intergenic region of these two clusters.

Example 2

Using PCR amplification techniques as described in USSN 08/587,680, Xa21 genes were isolated from cassava (SEQ ID NOS: 13-14), maize (SEQ ID NO: 15-16) and tomato (SEQ ID. NOs: 17-29). The following is a description of the methods used to isolate TRK1-7 from tomato. The same general procedure was used for maize and cassava.

We designed primers in conserved regions of both the Leucine Rich Repeat (LRR) region and the serine-threonine kinase domain of Xa21. The PCR products should amplify between these two domains and therefore span the transmembrane domain. So far, two sets of primers have proven successful to amplify three homologues of Xa21 in tomato.

The first clone TRK1 is a cDNA and the encoded polypeptide (SEQ ID NOs:17and 18). This clone is present as one or two copies in the tomato genome and one copy maps to the short arm of chromosome 1 in the proximity of a resistance gene to Xanthomonas campestris pv. vesicatoria (Rx1)(Zu *et al.* (1995) *Genetics* 41:675-682).

The second clone TRK2 (SEQ ID NO:19) is a 496bp PCR product with an ORF encoding a polypeptide (SEQ ID NO:20). TRK2 maps within a few cM of mcn (figure 4) a mutation on chromosome 3 that mimics disease lesions. A third clone TRK3

(SEQ ID Nos: 21 and 22) is a 473bp fragment and maps to chromosome 8 near an erecta like mutant. TRK4-7 (SEQ ID Nos: 23-29) are further PCR products and encoded polypeptides

Primers that have been proven useful are as follows.

1. LRR region

L3a. TCA AGC AAC AAT TTG TCA GGN CA(A/G) AT(A/C/T) CC

2. Kinase region

K1a CGC CTT AGG ATT TTC AAG CTT TCC (T/C)TT (G/A)TA NAC

K2a. TAA CAG CAC ATT GCT TGA TTT NAN (G/A)TC NCG (G/A)TG

K2b. TAA CAG CAC ATT GCT TGA TTT NAN (G/A)TC (G/A)CA (G/A)TG

K2c. TAA CAG CAC ATT GCT TGA TTT NAN (G/A)TC (T/C)CT (G/A)TG

The following combinations of primers are preferred:

L3a+K1a then L3u+K1u

L3a+K2a then L3u+K2u

L3a+K2b then L3u+K2u or

L3a+K2c then L3u+K2u.

PCR conditions

first cycle

94 for 30 s

55 for 30 s

72 for 1 min

For the next 19 cycles, the annealing temperature drops 1degree C every cycle. After 20 cyles, 10 min at 72. After inital amplification as second round of amplification is performed with the following specific primers with 1 microliter of the previous PCR.

L3u. TCA AGC AAC AAT TTG TCA

K1u. CGC CTT AGG ATT TTC AAG CTT

K2u. TAA CAG CAC ATT GCT TGA

The conditions for this amplification are:

35 cycles

94 15 sec

55 15 s

72 1 mn

after 35 cyles, 72 for 10 min

The above examples are provided to illustrate the invention but not to limit its scope. Other variants of the invention will be readily apparent to one of ordinary skill in the art and are encompassed by the appended claims. All publications, patents, and patent applications cited herein are hereby incorporated by reference.

WHAT IS CLAIMED IS:

1. An isolated nucleic acid construct comprising an *RRK* polynucleotide sequence, which polynucleotide hybridizes to SEQ ID NOs: 1, 4, 6, 8, 10, or 11 under stringent conditions.

2. The nucleic acid construct of claim 1, wherein the *RRK* polynucleotide sequence encodes an *RRK* polypeptide having an leucine rich repeat motif.

3. The nucleic acid construct of claim 1, wherein the *RRK* polynucleotide sequence encodes an *RRK* polypeptide having a cytoplasmic protein kinase domain.

4. The nucleic acid construct of claim 1, wherein the polynucleotide sequence is a full length gene.

5. The nucleic acid construct of claim 1, wherein the *Xa21* polynucleotide is as shown in SEQ ID NOs: 1, 4, 6, 8, 10, or 11.

6. The nucleic acid construct of claim 1, further comprising a promoter operably linked to the *RRK* polynucleotide sequence.

7. The nucleic acid construct of claim 1, wherein the promoter is a tissue-specific promoter.

8. The nucleic acid construct of claim 1, wherein the promoter is a constitutive promoter.

9. An isolated nucleic acid construct comprising a cassava *RRK* polynucleotide sequence, which polynucleotide hybridizes to SEQ ID NO: 13 under stringent conditions.

10. The isolated nucleic acid construct of claim 9, which is SEQ ID NO:

13.

11. An isolated nucleic acid construct comprising a maize *RRK* polynucleotide sequence, which polynucleotide hybridizes to SEQ ID NOs: 15 or 16 under stringent conditions.

12. The isolated nucleic acid construct of claim 11, which is SEQ ID NO: 15 or SEQ ID NO: 16.

13. An isolated nucleic acid construct comprising a tomato *RRK* polynucleotide sequence, which polynucleotide hybridizes to SEQ ID NOs: 17, 19, or 21 under stringent conditions.

14. The isolated nucleic acid construct of claim 13, which is SEQ ID NO: 17, SEQ ID NO:19, or SEQ ID NO:21.

15. A transgenic plant comprising a recombinant expression cassette comprising a plant promoter operably linked to a *Xa21* polynucleotide sequence of claim 1.

16. A method of enhancing resistance to *Xanthomonas* in a plant, the method comprising introducing into the plant a recombinant expression cassette comprising a plant promoter operably linked to an *RRK* polynucleotide sequence of claim 1.

17. The method of claim 16, wherein the plant tissue is from rice.

18. The method of claim 16, wherein the plant tissue is from tomato.

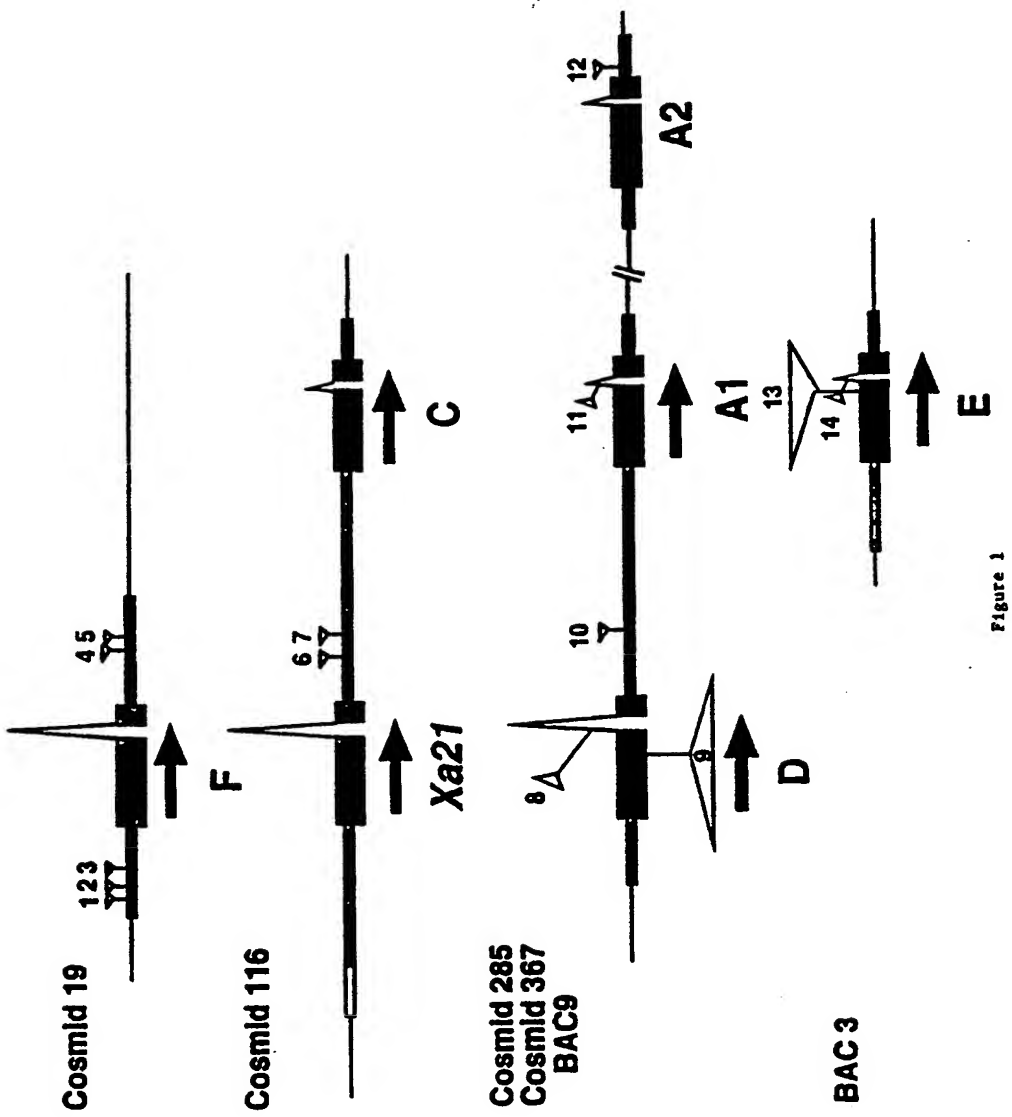


Figure 1

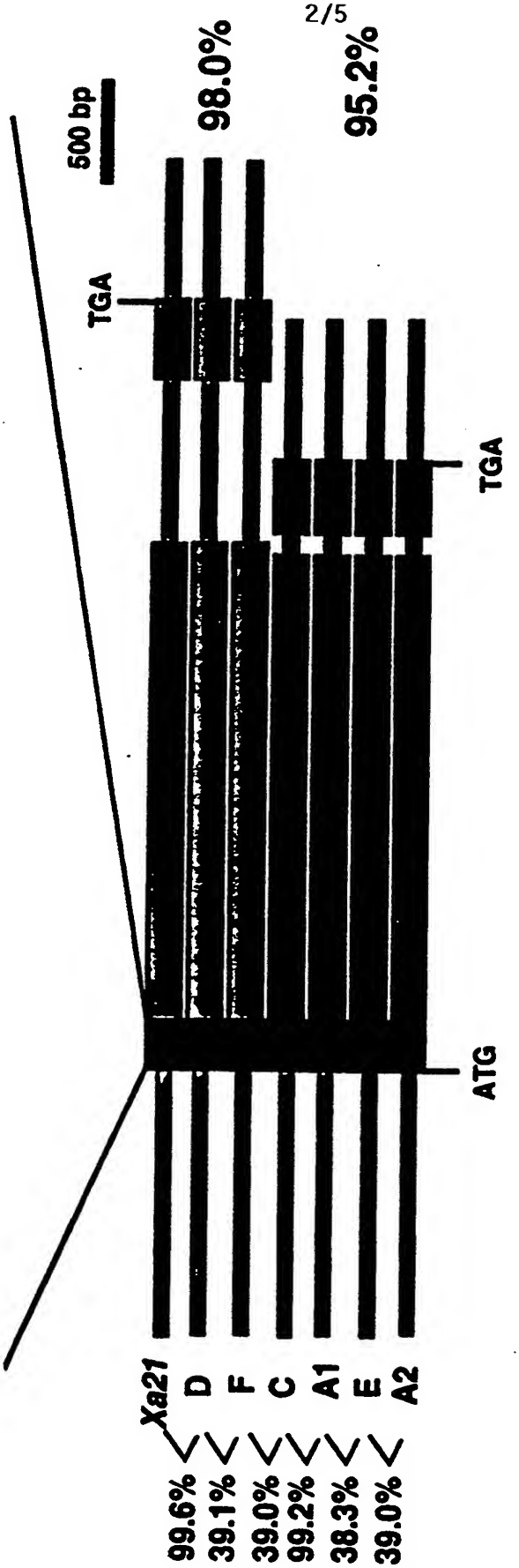
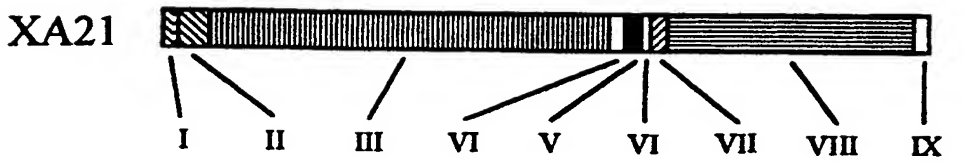


Figure 2A



A1(%)	100	100	59.5	57.1	34.6	83.3	64.0	82.0	38.5

Figure 2B

FIGURE 3A

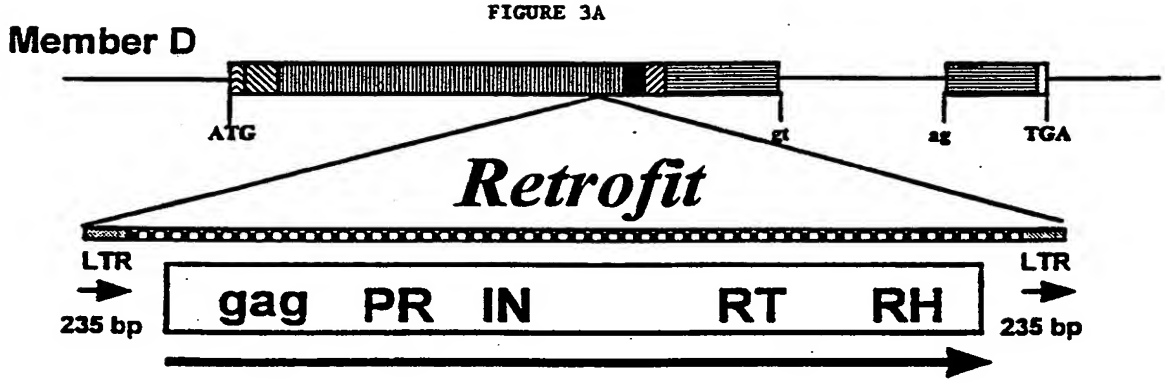
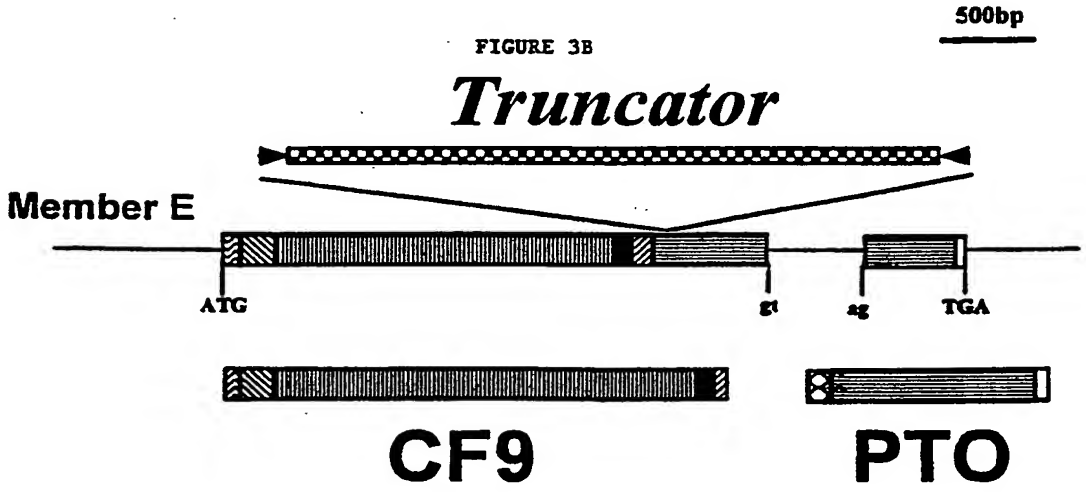


FIGURE 3B



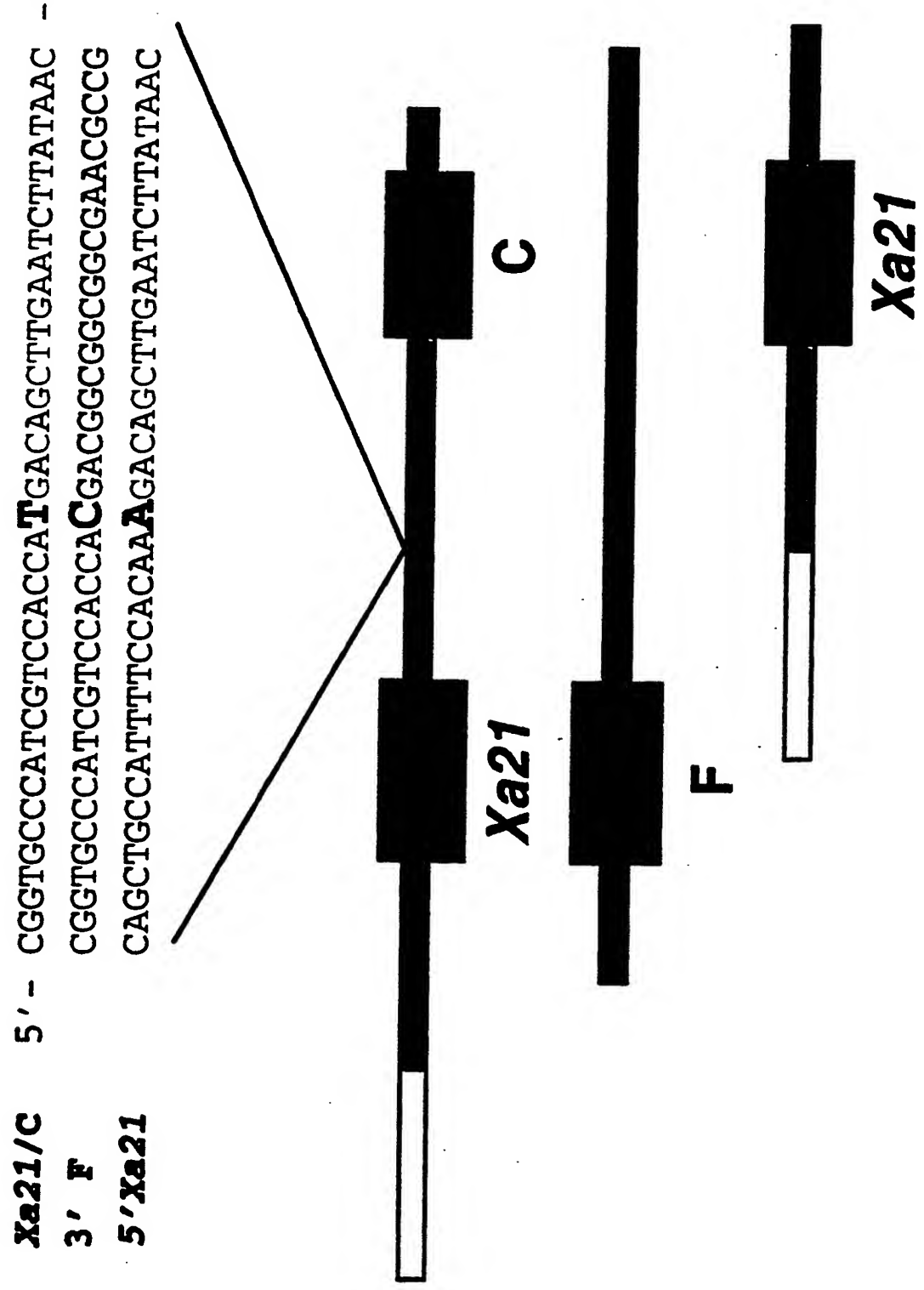


Figure 4

Sequence Listing

SEQ ID NO:1

```
DEFINITION  Oryza longistaminata receptor-like protein, family member D, and
            retrofit (gag/pol) genes, complete cds.
ACCESSION   U72726
SOURCE      long-staminate rice.
ORGANISM    Oryza longistaminata
            Eukaryotae; mitochondrial eukaryotes; Viridiplantae;
            Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
            Liliopsida; Poales; Poaceae; Oryza.
COMMENT     U72725 and U72726 are separated by a large AT rich
microsatellite region.
FEATURES             Location/Qualifiers
     source          1..13341
                     /organism="Oryza longistaminata"
                     /strain="IRBB21"
                     /chromosome="11"
                     /map="11q, RG103"
     CDS             2367..4205
                     /note="Xa21 gene family member D
                     /codon_start=1
                     /product="receptor kinase-like protein"

    misc_feature     4201..9071
                     /note="retrofit, a copia-like, transposon-like element"
     gene            4484..8821
                     /gene="gag/pol"
     CDS             4484..8821
                     /gene="gag/pol"
                     /codon_start=1
                     /product="retrofit"

     intron          9915..11712
     misc_feature    10020..10975
                     /note="Krispie, transposon-like element"
     3' flanking     12114..13341
     misc_feature    12626..12750
                     /note="Pop-012, transposon-like element"
                     13040..13248
                     /note="Ds-rice2, transposon-like element"
```

```
   1 aagcttcatt ggtttcttca gttatactta cgtaggtttt tcctgtatac ataaatacgt
  61 aacagagtaa gggaattaga ttgtttaaaa taaaatacat ataatctaat agcctaaaat
 121 atcaggtcca ctgacagtgg cggatctagg atttagaata tgggtggtcc gacctaattt
 181 tttcctaaac atactaaatc taacgatggt aatatatact atgcaagtat agataataga
 241 atagaccaaa agtgtatcat gctatattaa taaagcatct taaaacatat ataattaata
 301 attacctaaa attttgactt aaagaagctc acatggctat aaaagtttaa agaaaattac
 361 catactaatt tttcttctta tcgggtctac gccttctaat ggccatgaaa gtggtcgtta
 421 tatcttcttc cttcactctt aagaaaacat cccgcttaat ggatgtgtct atactatcat
 481 ccaaaagctc atcacccatc ttttttctca accatcatta gtaaatgcat cagttctact
 541 ataatttaat atcacaatgc acaggagtaa agagttcaaa atttcaaaac tgaaaattga
 601 aaaaaaagta aaaaaaaaat agaaaacctt tttgttttgg cttggtgcag gtctgcacca
 661 gtgcgctagt gcggcactgc ggcggcagcg gccaaggtgt cgacgcgcgt gcgtggcccg
 721 gtggcgctcg ctctcacgat ctgatcagat cgctgatcgc gtcggcgtcg cgactcgcga
 781 gggcgaggag gagagcgaca gagagtcctg cgacggcgcg acgcttcggt ttcttaattc
 841 cgaacgatta gatacaccgt acacgcgcgt gtggtgtggg gcctgtggta atctaatggt
```

```
 901 ttaaaatatt gggtccacca atttaagtga aaatcgacgg ttagatatga tagagctacg
 961 tggcagccta agagcgtttg taggagtccc acgtggcggt ttgagagcgt ttgtaggaag
1021 tttaatggac ttttagtata taatagatat ttataatttt attaagtacc ctaattttcc
1081 ctaaacaatt tttctctctc atcgtatttc catatatctt tttgagataa taatggatat
1141 aaacatagct agaaatgtaa atgttcacct tgcatcaata ggggatgaag ttgctaacct
1201 tttagatctc ctcgatttgt ataatataac caaaatattt tcaccaaaaa tttcgttaaa
1261 catccgagat atttgttgtt tttgccgatc gagcaaagat tagtagtcca gcagtgtctg
1321 caccaccacc atcgtgataa tgcatcttgt gtgttattct tgatgagaaa atacgtagtg
1381 aaaaccacat atgtggtgga aacttagaaa ctaccgttag atcgagaaat ggatgtccaa
1441 gattcgtcca cgtcaccaag agataaaatt taactcgcag attcacttat gagttaaaat
1501 tttaatgaga gttaaatttt aactcatgtt gatgtggacg aatatcggac atccatttct
1561 cgatccaacg atagcttcca agtttccact acatatgtgg tttgcactat atattttccc
1621 attcttgatt atgtgtttga gagcagctag cacaaagaga aaaaaaagca tcgtttttca
1681 cgcgtatgtt ttcagaactg ttaaatggtg tgtttttga aaaaactttc tatagaaaag
1741 tttctttaaa aaatatatta atctattttt taagtttaaa ataattacta cttaattaat
1801 tatacactaa cagcttattt cgttctacgt atcttgtcaa ttttcgctat tccttttcttc
1861 tcaaacacgg cattggatgc tctcatagca cttgctcgtt cggatagaag acttgacgaa
1921 gacgaccgct acaacttggt gtgttatatc gtgctttgtt tagcataatc attacatata
1981 ttccatgccg aagtgccgac gatgagaccg tgttcgatgc atctttgtat ggcatctagg
2041 gacaaagagc atagagtccc taccatagta cctgctcgcg tagaagactt gacgagaaga
2101 ccgactgcta caccttggtg tgtaataata tcgtgttgtg tgtaccatgc atactccttt
2161 aaaacaaata atggtggtaa cagtaaatct gtcatcccac ccactctcat tgtaaatttt
2221 gcaagttatc acttgaactt cttaatactc catccgtttg cgtgtgttct ttcagaattt
2281 gcgtgagcac ttttttcttct atataatctg tctagtccat gagctaaacc agcatctctc
2341 gctgtcttgc cttgcacttc tgcacgatga tatcactccc attattgctc ttcgtcctgt
2401 tgttctctgc gctgctgctc tgcccttcaa gcagtgacga cgatggtgat gctgccggcg
2461 acgaactcgc gctgctctct ttcaagtcat ccctgctata ccaggggggc cagtcgctgg
2521 catcttggaa cacgtccggc cacggccagc actgcacatg ggtgggtgtt gtgtgcggcc
2581 gccgccgccg ccggcaccca cacagggtgg tgaagctgct gctgcgctcc tccaacctgt
2641 ccgggatcat ctcgccgtcg ctcggcaacc tgtccttcct cagggagctg gacctcggcg
2701 acaactactt ctccggcgag ataccaccgg agctctgccg tctcagcagg cttcagctgc
2761 tggagctgag cgataactcc atccaaggga gcatccccgc ggccattgga gcatgcacca
2821 agttgacatc gctagacctc agccacaacc aactgcgagg tatgatccca cgtgagattg
2881 gtgccagctt gaaacatctc tcgaatttgt accttcacaa aaatggtttg tcaggagaga
2941 ttccatccgc tttgggcaat ctcactagcc tccaggagtt tgatttgagc ttcaacagat
3001 tatcaggagc tataccttca tcactggggc agctcagcag tctattgaat atgaatttgg
3061 gacagaacaa tctaagtggg atgatcccca attctatctg gaacctttcg tctctaagag
3121 cgttttgtgt cagcgaaaac aagctaggtg gtatgatccc tacaaatgca ttcaaaaccc
3181 ttcacctcct cgaggtgata tatatgggca ctaaccgttt ccatggcaaa atccctgcct
3241 cagttgctaa tgcttctcat ctgacacggc ttcagattga tggcaacttg ttcagtggaa
3301 ttatcacctc ggggtttgga aggttaagaa atctcacaga actgtatctc tggagaaatt
3361 tgtttcaaac tagagaacaa gaagattggg ggttcatttc tgacctaaca aattgctcca
3421 aattacaaac attgaacttg ggagaaaata acctgggggg agttcttcct aattcgtttt
3481 ccaatctttc cacttcgctt agttttcttg cacttcattt gaataagatc acaggaagca
3541 ttccgaagga tattggcaat cttattggct tacaacatct ctatctctgc aacaacaatt
3601 tcagagggtc tcttccatca tcgttgggca ggcttaaaaa cttaggcatt ctactcgcct
3661 acgaaaacaa cttgagcggt tcgatcccgt tggccatagg aaatcttact gaacttaata
3721 tcttactgct cggcaccaac aaattcagtg gttggatacc atacacactc tcaaacctca
3781 caaacttgtt gtcattaggc ctttcaacta ataaccttag tggtccaata cccagtgaat
3841 tattcaatat tcaaacacta tcaataatga tcaatgtatc aaaaaataac ttggagggat
3901 caataccaca agaaataggg catctcaaaa atctagtaga atttcatgca gaatcgaata
3961 gattatcagg taaaatccct aacacgcttg gtgattgcca gctcttacgg catctttatc
4021 tgcaaaataa tttgttatct ggtagcatcc catcagcctt gggtcagctg aaaggtctcg
4081 aaactcttga tctctcaagc aacaatttgt caggccagat acccacatcc ttagcagata
4141 ttactatgct tcattccttg aacctttctt tcaacagctt tgtgggggaa gtgccaacca
4201 tgtgaaacag tctgttctat tatcctggac atgagatata catgcatgct ccagtacatg
4261 cgatgctcta gtactgggat atggttcgtt aggataagga gattatctcc aaatttgttt
4321 caccttctta tccattgtaa ctgaatactg taagccacgg tagaggctgt ttctaccact
```

```
4381 atttaacaga gatgcggccc aaggcaaagg gtttaacgct tcacatttta tcacatggta
4441 tcagagcctt tttccacctc aatagatcgc atctttcatc cacatggcgt cgtcgtcgtc
4501 ttcctcaggc gcggcagccg ccaatctcct ccaaggccac tcggtttcag agaaactcgg
4561 gaaagccaac catgcattgt ggaaagcgca agttagcgct gcagtgcgtg gagcccgatt
4621 gctgggctac ctcaacggcg atatcaaagc tccagacgcc gaactctcgg tcaccataga
4681 tgggaagacc acaacaaagc cgaatccggc atttgaagat tgggaggcca atgaccagct
4741 tgttcttggc tatctcctgt catctctttc aagggatgtg ctgatccaag tcgccacatg
4801 caagacggcg gctgaggcat ggcggagcat tgaagcactc tactccaccg gcactcgagc
4861 aagggcggtg aacaccagac tcgccctcac caacacgaag aaaggaacaa tgaagatcgc
4921 cgagtatgtc gccaagatgc gagcgcttgg tgatgagatg gctgccggcg gtcatccact
4981 tgatgaagaa gaccttgtcc agtacatcat cgctgggcta aatgaagact tcagcccgat
5041 cgtctccaac ctctgcaaca agtccgatcc catcacggtt ggggagctgt attctcagct
5101 cgtcaacttt gaaaccctcc ttgatctcta ccgcagcact ggtcagggag gagctgcttt
5161 tgtcgctaat cgcggcaggg gcggcggcgg cggcgggcgc ggcaacaaca acaactccgg
5221 cggcggcggc ggcagaagcg cgccgggtgg acgcggcagc ggcagccagg gtcgcggtgg
5281 ccgtggacgc ggcacaggag gccaagacag gcgccctact tgccaagttt gtttcaagcg
5341 tgggcataca gcagctgatt gttggtatcg cttcgacgag gactacgttg cagatgagaa
5401 gctcgttgct gctgctacta actcgtatgg tatagataca aattggtata ttgatacagg
5461 tgctacagat cacattaccg gtgaactaga gaagcttacc accaaggaga aatacaacgg
5521 cggcgagcaa attcacactg ctagcggagc aggtatggat attagtcaca ttggtcatac
5581 tattgtgcat acccctagcc gtaatattca tctaaataat gtcctttatg ttcctcaagc
5641 caagaaaaat cttatatctg ctagtcaatt agccgctgat aattctgctt ttcttgaact
5701 tcactcgaaa ttctttttcta taaaggatca ggtaacgagg gacgttctgc ttgaagggaa
5761 atgtagacac ggtctctacc cgatccccaa gttctttggt cgctcaacca acaaacaagc
5821 ccttggtgcc gccaagttat ccctgtctag gtggcatagc cgtctaggac atccgtctct
5881 tcctattgtt aagcaagtca ttagcagaaa taatctccca tgttcagttg agtcagtcaa
5941 tcagtctgtg tgtaatgctt gccaagaagc aaagagtcat cagttacctt atattagatc
6001 tactagtgtg tctcaatttc ctcttgaact tgtttttttct gatgtttggg gccctgctcc
6061 agagtctgtt gggagaaata aatattatgt gagtttcatt gatgatttta gtaagtttac
6121 ttggatatac ttgctgaaat acaagtctga ggtttttgag aaatttaaag aatttcaggc
6181 tttagttgaa cgaatgtttg atagaaagat tattgccatg cagactgatt ggcgggggggg
6241 gagatatcag aaacttaatt cctttttttgc tcaaatagga ttgatcatca tgtgtcatgt
6301 cctcacactc atcaggcaga atgggtcagc tgagagaaaa caccggcata tcgtggaagt
6361 aggcctttct cttttatctt atgcatcaat gcctcttaag ttttgggatg aagcctttgt
6421 tgcagccact tatctcatca atcgtatacc tagtaaaacc atccaaaatt ctacacccct
6481 agagaaactg tttaaccaaa aacctgacta ctcatccttg agagtgtttg gttgtgcatg
6541 ttggcctcat cttcgcccct acaatacaca caaactccag tttcgctcca aacagtgcgt
6601 gtttttgggt tttagtactc accacaaagg atttaagtgt cttgatgtgt catcaggccg
6661 tgtctacatc tcaagagatg ttgtctttga tgaaaatgtt tttcccttct ctacactcca
6721 ctcaaatgca ggagccagac tcaggtctga aattcttttg ttgccgtccc ccttgacaaa
6781 ctataatacg gctagtgcag ggggaacaca tgtagttgca ccagtggcta atactccatt
6841 acctagtgat aatttaattt ctaatgctgc tgatgtgact tctggagaaa atagtgcagc
6901 acatgaacag gaaatggaga atgagcagga aatagagaac gtcatgcatg ggaacgacgt
6961 gcatggggac gcggcatcgg gacctgtgct ggatcaacca actgctgaca gcagcactgc
7021 gccggaccag ggagctgaca ccagtgacgc ggtctctggc gcagcttctg acgcgggtgg
7081 agacactgcc accctgggag ctggagcagc aaatagcgca gcagcaggtg gtgaagaatc
7141 ccagccggtg cagcctgatg tgacgggtac agtactggct acagtagccc ctgcatcgag
7201 accacacact cgtctgcgga gtggtattcg aaaagagaag gtatacactg atggcaccgt
7261 taaatatggt tgtttttctt ctactggtga accacaaaat gataaagagg ctttaggaga
7321 taaaaactgg agagatgcaa tggaaactga gtataatgct ttgataaaaa atgacacatg
7381 gcacctagtt ccatatgaga aaggacaaaa tatcattggg tgtaaatggg tatataagat
7441 taaaaggaag gcagatggga cacttgatag atacaaagct agacttgtag caaaggggtt
7501 taaacaaaga tatggtatcg attatgaaga tacttttagt cctgttgtta aagctgctac
7561 tattagaatt attctgtcca ttgctgtctc tagaggttgg agtcttagac agttagatgt
7621 tcagaatgcc tttcttcatg gattcttaga agaagaagtc tacatgcaac aacctcctgg
7681 gtttgagtca tcctctaaac ctgattatgt atgtaaattg gataaggcat tatatgggct
7741 gaaacaagca ccaagggcgt ggtattccag gctgagtaag aaacttgttg aacttggttt
7801 tgaagcttca aaggctgata cctcattatt ctttcttaac aaaggaggga tacttatgtt
```

```
 7861 tgttttggta tatgttgatg atataattgt agctagctct acagagaagg caactacagc
 7921 acttctgaag gatctaaaca aggagttcgc acttaaggat ttgggagacc tgcactactt
 7981 ccttggaatt gaggtaacta aagtttccaa tggcgttatc ttgactcaag agaagtatgc
 8041 aaatgatctg ctaaagagag ttaatatgtc aaattgcaag ccagttagta ctcctctttc
 8101 tgttagtgaa aaattaactc tatatgaggg atcacccttg ggtcctaatg atgcaataca
 8161 atatagaagt atagttggtg ctttacaata cttgaccttg acaagacctg acatagctta
 8221 ttcagtaaac aaagtctgtc agtttcttca tgctcctact accagtcatt ggattgcagt
 8281 aaaaagaatc ctcagatact tgaaccaatg cacaagtcta ggacttcata tacacaagag
 8341 tgcttctact cttgttcatg ggtattctga tgcagactgg gcaggtagta tagatgacag
 8401 aaaatcaaca ggaggatttg cagtattttt gggttctaat cttgtgtcct ggagtgctag
 8461 gaaacaacct actgtgtcaa ggtcaagcac agaggcagaa tataaggctg tggcaaatac
 8521 tacagccgaa ctgatatggg tacaaacctt gttaaaagaa ttgggaattg agtctcctaa
 8581 agctgccaag atttggtgtg ataacttagg agctaaatat ttatcagcta atcctgtgtt
 8641 tcatgcaagg acaaagcata tagaggttga ttatcatttt gtaagagaac gagtgtcaca
 8701 gaagctgtta gagattgatt ttgttccatc aggagaccaa gttgctgacg ggtttacaaa
 8761 ggcactgtca gcttgtcttc ttgaaaattt taaacacaat cttaacctag ctaggttatg
 8821 attgagaggg ctgtgaaaca gtctgttcta ttatcctgga catgagatat acgtgcatgc
 8881 tccagtacat gcgatgctcc agtactggga tatggttcgt taggataagg agattatctc
 8941 caaatttgtt tcaccttctt atccattgta actgaatact gtaagccacg gtagaggctg
 9001 tttctaccac tatttaacag agatgcggcc cgaggcaaag ggtttaacgc ttcacatttt
 9061 atcacaaacc attggtgctt tcgcagctgc atccgggatc tcaatccaag gcaatgccaa
 9121 actctgtggt ggaatacctg atctacatct gcctcgatgt tgtccattac tagagaacag
 9181 aaaacatttc ccagttctac ctatttctgt ttctctggtc gcagcactgg ccatcctctc
 9241 atcactctac ttgcttataa cctggcacaa gagaactaaa aagggagccc cttcaagaac
 9301 ttccatgaaa ggccacccat tggtctctta ttcgcagttg gtaaaagcaa cagatggttt
 9361 cgcgccgacc aatttgttgg gttctggatc atttggctca gtatacaaag gaaagcttaa
 9421 tatccaagat catgttgcag tgaaggtact aaagcttgaa aatcctaagg cgctcaagag
 9481 tttcactgcc gaatgtgaag cactacgaaa tatgcgacat cgaaatcttg tcaagatagt
 9541 tacaatttgc tcgagcattg ataacagagg gaacgatttc aaagcaattg tgtatgactt
 9601 catgcccaac ggcagtctgg aagattggat acaccctgaa acaaatgatc aagcagacca
 9661 gaggcacttg aatctgcatc gaagagtgac catactactt gatgttgcct gcgcactgga
 9721 ctatcttcac cgccatggcc ctgaacctgt tgtacactgt gatattaaat caagcaatgt
 9781 gctgttagat tctgatatgg tagcccatgt tggagatttt gggcttgcaa gaatacttgt
 9841 tgatgggacc tcattgatac aacagtcaac aagctcgatg ggatttagag ggacaattgg
 9901 ctatgcagca ccaggtcagc aagtccttcc agtattttgc attttctgat ctctagtgct
 9961 atatgaaata gtttttacct ctagtgaaac tgatggagaa tataagtaat taattgaact
10021 caggggcgaa gcagaataaa tcatcgccgg ggtcactact aactaatgaa cttgcactac
10081 tatgaacagg tttatcgtaa atggccaaaa cacattttta caggtggcca atcgtgtcgc
10141 ttatgactaa agatcttttc gcttgtgacc ggtgctacga taatcacatg gtgaaatgaa
10201 tcttcacaga tatagccttt aagtttaagc tagccatttg cagaaaatga aaggggtggc
10261 tgtcagaaaa aaaaaaccat ttttttgtag tgtgactaaa actatgcgta agatggaaca
10321 aaattataca ttaatcctcc tccttactca tgatcatata actgaagttt gaaaacaaag
10381 acaaatcaaa tgttcatatt tcagaacttc tgaattatag aaccctaact tcctaataaa
10441 ctttcggact tgagaaacca ccctagggct aaattgtgat ggtaataatc aagaaattgt
10501 gataatgtta gtaggttacc gccctcaaca gctcgagcgc tcgagcgctc aaccataacc
10561 cttcaccctt gccattcttc tcgtccttgt tacgccaagc cggcgatgca cgcgtgtact
10621 actcgaggac gataagacgc ggaaagtcga gaaatcgaga tgtgcggcga tgcaatgcga
10681 agtgctcggt caaactacca gaagaatatg atgcaacagg ggaatggact ttctgggctg
10741 ccggccatgt gggctcaagt ctcaggcgtt taagagctga ttaattttct attttctact
10801 tactagtctc gttcctaatt gctattgggc tttaacaaaa cggtgtcacc ggggtccagt
10861 agtttgcttg actggggtat agctagtaaa aaagcgaggg gtactagatg tatgcacgaa
10921 atcgtcgggg tctactgacc ccgatggtag gctgtagcgt cgcccctgat tgaactaatt
10981 aaattgcaca aaaataagat tatttgccat atctattcag atgctaaata tagctagttc
11041 atagaggtac agatttttt atataggact ctagagctac cacacactca aatcaaatta
11101 tgggtgtgtt ctgctctaca ctgcaatatg aaatgattat tacttctaca tgaactgatg
11161 gaggagtttc agaaggatca aatttgagta aaattttcaa ttctacattt aagaaacact
11221 tttttttcat atgctagtta caatttttta tttcacgagc ttacattgac catgaaaaat
11281 acttggcact acttactaat tcccacatgg aggtagtgaa aataatatag atacaaaaac
```

```
11341 gaaatatcct atgttgtgtg atatactata atcacaatga acacaaacag gattcgtaca
11401 aaagtaatta gccatcatag caactgattg cttggggtaa ctgtatagca caatcatacc
11461 aaatttcttt agatatgtat ctgtaaatta gattcttaaa gttaaatatg aaatttcatt
11521 ggtatttatg tttctttata taataaaaat taatccagcc tttgcatcta tcatttgtcc
11581 agacatcctt gttatttgtg atatttaaca cgtaaattta cataattata catccaagtt
11641 ctttttattt aacactgtaa atttcaaatc gtacatgtta taaagaatgt actatatttc
11701 ctgctcaaac agagtatggc gttgggctca ttgcatcaac gcatggagat atttacagct
11761 atggaattct agtgctggaa atagtaaccg ggaagcggcc aactgacagt acattcagac
11821 ccgatttggg cctccgtcaa tacgttgaac tgggcctaca tggcagagtg acggatgttg
11881 ttgacacgaa gctcattttg gattctgaga actggctgaa cagtacaaat aattctccat
11941 gtagaagaat cactgaatgc attgtttggc tgcttagact tgggttgtct tgctctcagg
12001 aattgccatc gagtagaacg ccaaccggag atatcatcga cgaactgaat gccatcaaac
12061 agaatctctc cggattgttt ccagtgtgtg aaggtgggag ccttgaattc tgatgttatg
12121 tctcgtaatg ttttattgcc actcttcaga tcgacttctg cagtggtatc taccacacga
12181 tcactaaagt caccgtggtt atttcctgat ccagcatatc tgatcatgca tgttctgtgt
12241 tgtatacctg tattttactc tgaattgcca caccgcaacc ctgcatctgt ttgtttggta
12301 tacaaaagat agtgatgagt ttattgtttt aggggcttcc tagttggcgc gtgtggtgcc
12361 ggcacgcacg cagcccgagg gtgggtttct ttttttcca ttgttattcc gttgcttttt
12421 ttcaccacgg tagatctttt ttttccggat ttccattttt tccgttgttt ttctctatcg
12481 cttatgttgg cggatttttt tccgtggttt tctttccgaa gacgagtata tctaacgtaa
12541 ctaacatgtt acttttagat aacgatggtt attaagataa gatttttctc tggaagattt
12601 ttgtaagtaa cagattgaaa acaaatctat acgtgaggtc aaattttgaa aactttcaat
12661 ctagatttaa aaacttttca actcaaaatt tgaatttttg aagtgaaaat ttgaaaactt
12721 tcaaaaatta ctagtaatcg acaaaaaaaa aatggaaatg gaaacggaaa tagttttgct
12781 gttataccga tcgtttccat atttaccgta ttcttataga aattaccgtt tcttataata
12841 tggtaattac cgtatttcta aatatgttga tatttatagg gcatgtctct acttgactca
12901 cagtttagag attgattgac tatttaatca aatccctaac ttgattgcat ggctaaaatg
12961 gagttgattt ctaatttata tagtatagct tgaatttatt tgtaaatata acatacttat
13021 gtaaagttaa atatatgttt tctatagttt aatgtttctg tatttgttac cggttttcga
13081 tctgtaccga catgtttcca tcagtattat tccatgtccg gttttccgat atttccgata
13141 tcgttttcgt ttccgacttt accgttttcg atttcatttc cgagaaaaat atgattatgg
13201 aaatggtcga ggctgttttc cgatcgtttc cgaccgtttt catctctacc cgtagtaata
13261 atatataaca ttttatctct aatctttctc tctctcatat caatgaaata atcgctaaga
13321 gactgctatt aacaagggct t
```

SEQ ID NO:2

/translation="MISLPLLLFVLLFSALLLCPSSSDDDGDAAGDELALLSFKSSLL
YQGGQSLASWNTSGHGQHCTWVGVVCGRRRRRHPHRVVKLLLRSSNLSGIISPSLGNL
SFLRELDLGDNYFSGEIPPELCRLSRLQLLELSDNSIQGSIPAAIGACTKLTSLDLSH
NQLRGMIPREIGASLKHLSNLYLHKNGLSGEIPSALGNLTSLQEFDLSFNRLSGAIPS
SLGQLSSLLNMNLGQNNLSGMIPNSIWNLSSLRAFCVSENKLGGMIPTNAFKTLHLLE
VIYMGTNRFHGKIPASVANASHLTRLQIDGNLFSGIITSGFGRLRNLTELYLWRNLFQ
TREQEDWGFISDLTNCSKLQTLNLGENNLGGVLPNSFSNLSTSLSFLALHLNKITGSI
PKDIGNLIGLQHLYLCNNNFRGSLPSSLGRLKNLGILLAYENNLSGSIPLAIGNLTEL
NILLLGTNKFSGWIPYTLSNLTNLLSLGLSTNNLSGPIPSELFNIQTLSIMINVSKNN
LEGSIPQEIGHLKNLVEFHAESNRLSGKIPNTLGDCQLLRHLYLQNNLLSGSIPSALG
QLKGLETLDLSSNNLSGQIPTSLADITMLHSLNLSFNSFVGEVPTM"

SEQ ID NO:3

/translation="MASSSSSSGAAAANLLQGHSVSEKLGKANHALWKAQVSAAVRGA
RLLGYLNGDIKAPDAELSVTIDGKTTTKPNPAFEDWEANDQLVLGYLLSSLSRDVLIQ
VATCKTAAEAWRSIEALYSTGTRARAVNTRLALTNTKKGTMKIAEYVAKMRALGDEMA
AGGHPLDEEDLVQYIIAGLNEDFSPIVSNLCNKSDPITVGELYSQLVNFETLLDLYRS
TGQGGAAFVANRGRGGGGGGRGNNNNSGGGGGRSAPGGRGSGSQGRGGRGRGTGGQDR
RPTCQVCFKRGHTAADCWYRFDEDYVADEKLVAAATNSYGIDTNWYIDTGATDHITGE
LEKLTTKEKYNGGEQIHTASGAGMDISHIGHTIVHTPSRNIHLNNVLYVPQAKKNLIS
ASQLAADNSAFLELHSKFFSIKDQVTRDVLLEGKCRHGLYPIPKFFGRSTNKQALGAA
KLSLSRWHSRLGHPSLPIVKQVISRNNLPCSVESVNQSVCNACQEAKSHQLPYIRSTS
VSQFPLELVFSDVWGPAPESVGRNKYYVSFIDDFSKFTWIYLLKYKSEVFEKFKEFQA
LVERMFDRKIIAMQTDWRGGRYQKLNSFFAQIGLIIMCHVLTLIRQNGSAERKHRHIV
EVGLSLLSYASMPLKFWDEAFVAATYLINRIPSKTIQNSTPLEKLFNQKPDYSSLRVF
GCACWPHLRPYNTHKLQFRSKQCVFLGFSTHHKGFKCLDVSSGRVYISRDVVFDENVF
PFSTLHSNAGARLRSEILLLPSPLTNYNTASAGGTHVVAPVANTPLPSDNLISNAADV
TSGENSAAHEQEMENEQEIENVMHGNDVHGDAASGPVLDQPTADSSTAPDQGADTSDA
VSGAASDAGGDTATLGAGAANSAAAGGEESQPVQPDVTGTVLATVAPASRPHTRLRSG
IRKEKVYTDGTVKYGCFSSTGEPQNDKEALGDKNWRDAMETEYNALIKNDTWHLVPYE
KGQNIIGCKWVYKIKRKADGTLDRYKARLVAKGFKQRYGIDYEDTFSPVVKAATIRII
LSIAVSRGWSLRQLDVQNAFLHGFLEEEVYMQQPPGFESSSKPDYVCKLDKALYGLKQ
APRAWYSRLSKKLVELGFEASKADTSLFFLNKGGILMFVLVYVDDIIVASSTEKATTA
LLKDLNKEFALKDLGDLHYFLGIEVTKVSNGVILTQEKYANDLLKRVNMSNCKPVSTP
LSVSEKLTLYEGSPLGPNDAIQYRSIVGALQYLTLTRPDIAYSVNKVCQFLHAPTTSH
WIAVKRILRYLNQCTSLGLHIHKSASTLVHGYSDADWAGSIDDRKSTGGFAVFLGSNL
VSWSARKQPTVSRSSTEAEYKAVANTTAELIWVQTLLKELGIESPKAAKIWCDNLGAK
YLSANPVFHARTKHIEVDYHFVRERVSQKLLEIDFVPSGDQVADGFTKALSACLLENF
KHNLNLARL"

SEQ ID NO:4

```
DEFINITION  Oryza longistaminata receptor kinase-like protein gene, family
            member A1, complete cds.
ACCESSION   U72725
SOURCE      long-staminate rice.
  ORGANISM  Oryza longistaminata
            Eukaryotae; mitochondrial eukaryotes; Viridiplantae;
            Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
            Liliopsida; Poales; Poaceae; Oryza.
COMMENT     U72725 and U72726 are separated by a large AT rich
microsatellite region.
FEATURES             Location/Qualifiers
     source          1..8416
                     /organism="Oryza longistaminata"
                     /strain="IRBB21"
                     /chromosome="11"
                     /map="11q, RG103"
     CDS             join(4771..7384,7676..8052)
                     /note="Xa21 gene family member A1; downstream of
                     microsatellite region; disease resistance gene family
                     member"
                     /codon_start=1
                     /product="receptor kinase-like protein"

     misc_feature    7432..7614
                     /note="Snap-O11, transposon-like element"

                     BASE COUNT     2220 a   1984 c   1707 g   2505 t
ORIGIN
        1 gccgtcgatc tgtcattttg aaacggccca ttcttttcca tctatatgca ttcatgaaat
       61 acatggtata tcccatcgat cggacatcac ctgttagcgc gtacgccatc gtcgtcatca
      121 acctagctag ggcaaacgca ccttactgag ctccgtcctc cgatcgccac catcaccaat
      181 gaacaagctg ctgcggcctc tcggtggcct gaggttgctc aaccgagaag aacatccgtt
      241 ccgatgcttc tcctcctcca tcgatctcgt cttcccaggt cgccgccgcc gccacatggc
      301 aaccaccgtg acccacccgc cgccaacgga atccgctggt tcgacggcgg cggccgcgac
      361 tgctgacccg ggctcggtga tgctggaacg ttggggctgc ctcaggggct ccacggcggc
      421 gaacgtagtc gccgacgaca acaccgcggg agttccgcac cttccgcggc caacccctcc
      481 ggtcgcctcg tccgggttgc gccgggatct ccttcatctg cttcgatcgc ggggttgatg
      541 gctacgtcat cgcggctcac ggcgactctg tcctcttccg gatgagttgg aacgactact
      601 tcgtctacat ggccgccgcc ggcaagccgc cgtcgctgac gctgctcccc gtctgcgaca
      661 tccccatgaa cgagcgctgc tgggtcagca aggaccgttt caaggacgct tccgcaccac
      721 gggccgggtg ttcgaccagc aggacaccgg catcctacgc ctccgcggcg acgacggcgg
      781 tgaggaggcg ccgctctagt ggcgcagtcc agatcgcgca cgagccgccg ttcgacacgg
      841 ccgagctctg cgtgctccgc cccggccacg gcgagtggga gctcaagatg gcggtgccca
      901 tcgtccacca tgacagcttg aatcttataa caaaccttgc tgaagctgac aaattctagc
      961 ccccagccat gaagttggaa aaatcaattt ccgattacac aaattggtta atacgcaacc
     1021 atttagtgct cttaacatga ccaggtttta catgttcgtt cggctcttag aatctgacaa
     1081 gaccttatct gctctgggcg tccccagccg aaattccatt agttttctcg gaggcttgtc
     1141 agaacagcgt aaagggacaa taggactgcc ttcaagatga ggcgatataa gaggggatca
     1201 acagacaaat attgcacata taaaacttac agaagttgat gtagatgatg agacgaccac
     1261 cacactaggc aaagaccagg tgtatagttg tactcaacaa atcacagagg tagtgagaga
     1321 tcgctacgat ctactggtcg aagatcaggt gtaggcgtat tctcgatcac ctgaagaaga
     1381 atctttaggt gttgagagat cgctactatc tactggtcaa tactagtaaa aaaacctcat
     1441 agagatcggc actataggtg ccggaacagc taaaaccggc acctataata cttttccctc
     1501 ctccgtggac tcaaagcacg taaaaccgac acctttaagc aactataggt gccggttcta
     1561 aagaagaacc gacacctata gtataggtgc tggtttttta aaaaaacccg acacctttaa
     1621 tataatatag gtgtcggttc ttctttaaaa ccgacaccaa taataaatta tacgtgtcgg
     1681 ttttttaata aaaccggcac ctatccaaac cgagcctagc tgtcgagtcg agccaatcca
     1741 ggctgacgca tatattagtc tcgtccttat cgcctcgtct ctctctctct ctctctctct
```

1801 ctcttctctc tcgcctctct gtgtagcgcg cggcggtggt cgggcggcgt cccggcgtag
1861 gcaacagtgg tggcgagatg ggaggcggtg gcgcatcaca gctattcact ttagcgcatt
1921 gaataataat cggcatcatg atctacttat ggttcgtgca agggggaaga ttgtagatac
1981 acatggtcac agctaagtgc tatgatcgcg gctcatctct ccaaatagat tcatgccatc
2041 cgtacttaaa cagaacctta tgttccgtgc atcctttctg aagtcttgaa atgttttatc
2101 gatgttttac cactacgacc catctgtagg gtgtctcgca tcccgtcttg tttatgtttt
2161 tatgtgtgcc agcgcaccat tctagcacgc gctttgttcc tgaacaaaca ccttagccgt
2221 ggtattaaat ggaaatactg catgatcata gcaggaactc tcttcttcga aggatgtccg
2281 tcaacatcac ctaggtcatc tcgtctaatc ttatatcgta gtgccttaca aaccaggcgt
2341 gcttctaggt tctcgtactc ctcaacgcga tagaggatac aatcattcag acatacgtgt
2401 atcttctgta cttccagtcc gagagggcag attacctttt tagcttcgta cgttgtttcg
2461 agcaattcgt ttccctcggg aagaatattc tttacgagtt tcaataactc gccaaatgcc
2521 ttgtcagtca caccatattt tgccttccat tgtcagaatt ccagtgtggt acccaacttt
2581 ttgtgctcct attcgcaacc tgggtacaag gacgttttgt ggtatgctaa catccggtcc
2641 aatttctaga ccaacttttt cacttttgca gtcatttttt acatcgcaca acatttgacc
2701 aagattatcc gtaccgccgt tttcttcaac agctctgtcc gcctcgccta ttgtattttc
2761 tgcaaatcca tcatattgag cccagtctga aatgttgtcg tcttcttatt aatttttcttc
2821 tattgcaacc cctaactctc cgtgcaaggt ctaacaatta tagccaggta tgaatcccga
2881 ttccaacaag tggatatgaa gacttcttag atgtagaata ctccttctgg tttttgcact
2941 tttgcatgga caacatataa aacccttagg cttgtgggca ttggccacac tcaaaaaata
3001 atgcacgtca tcaataaact ccttcgaccg tcattctgca gtgtacatcc attaccgatt
3061 catctacatt aggagataat aattgtaaag gaagtccaca aaagtgaaac tacttaaata
3121 atcatataat aatttaaaat atcataatta aattgaaaac tgacggtttt aatgtattct
3181 tcttatttct aacgatttta acgagtttta aatggactta atcggagtca tgattaacta
3241 tttataaatt ttatctgtct caataaattg taaatatatt tttccatgta ttatccttgt
3301 ttaattattt ttaaaagttc taaacatatt tttaatgcat tctacttatt tctaactatt
3361 ttaaagattt tcaaatggac ttagttttct atttttattc ttcatttttct atatttgccc
3421 tttgttgtct cttttaaca attttataca aatatttata attttattaa gtaccctaat
3481 tttccctaaa caatttttct ctctcatcgt atttccatat atctttttga gataataatg
3541 gatataaaca tagctagaaa tgtaaatgtt caccttgcat caatagggga tgaagttgct
3601 aacctttttag atctcctcga tttgtataat ataaccaaaa tattttcacc aaaaatttcg
3661 ttaaacatcc gagatatttg ttgtttttgc cgatcgagca aagattagta gtccagcagt
3721 gtctgcacca ccaccatcgt gataatgcat cttgtgtgtt attcttgatg agaaaatacg
3781 tagtgaaaac cacatatatg gtggaaactt ggaaactacc gttagatcga gaaatggatg
3841 tccaagatcg tccacgtcac caagagataa aatttaactc gcagattcac ttatgagtta
3901 aaattttaat gagagttaaa ttttaactca tgttgatgtg gacgaatatc ggacatccat
3961 ttctcgatcc aacgatagct tctaagtttc cactacatat gtggtttgca ctatatattt
4021 tcccattctt gattatgtgt ttgagagcag ctagcacaaa gagaaaaaaa agcatcgttt
4081 ttcacgcgta tgttttcaga actgttaaat ggtgtgtttt tttaaaaaac tttatataga
4141 aaagtttctt taaaaaatat attaatctat tttttaagtt taaaataatt actacttaat
4201 taattataca ctaacagctt atttcgttct acgtatcttg tcaattttcg ctattccttt
4261 cttctcaaac acggcattgg atgctctcat agcacttgct cgttcggata gaagacttga
4321 cgaagacgac cgctacaact tggtgtgtta tatcgtgctt tgtttagcat aatcattaca
4381 tatattccat gccgaagtgc cgacgatgag accgtgttcg atgcatcttt gtatggcatc
4441 tagggacaaa gagcatagag tccctaccat agtacctgct tgcgcagaag acttgacgag
4501 aagaccgact gctacacctt ggtgtgtaat aatatcgtgt tgtgtgtacc atgcatactc
4561 ctttaaaaca aataatggtg gtaacagtaa atctgtcatc ccacccactc tcattgtaaa
4621 ttttgcaagt tatcacttga acttcttaat actccatccg tttgcgtgtg ttctttcaga
4681 atttgcgtga gcacttttc ttctatataa tctgtctagt ccatgagcta aaccaacatc
4741 tctcgctgtc ttgccttgca cttctgcacg atgatatcac tcccattatt gctcttcgtc
4801 ctgttgttct ctgcgctgct gctctgccct tcaagcagtg acgacgatgg tgatgctgcc
4861 ggcgacgaac tcgcgctgct ctctttcaag tcatccctgc tataccaggg gggccagtcg
4921 ctggcatctt ggaacacgtc cggccacggc cagcactgca catgggtggg tgttgtgtgc
4981 ggccgccggc acccgcacag ggtggtgaag ctgcggctgc gctcgtccaa cctgaccggg
5041 atcatctcgc catcgctggg caacctatcc ttcctcagga cgctgcaact cagcaacaac
5101 cacctgtccg gcaagatacc ccaggagctc agccgtctca gcaggctcca gcagctggta
5161 ctgaatttca acagcctatc gggtgagatt ccagctgctt tgggcaatct aaccagtctc
5221 tcagttcttg agctgactaa caatacactg tctggttcta tcccttcatc cctgggcaag

```
5281 ctcaccggcc tctataatct tgcactggct gaaaatatgc tgtctggttc catccctacg
5341 tctttcggcc aattgcgcag attatctttc cttagcttag ccttcaacca cttaagtgga
5401 gcgatcccag atcctatttg gaacatctcc tctctcacca tatttgaagt cgtgtccaac
5461 aacctaactg gtacactgcc tgcaaatgca ttcagtaatc ttcctaatct gcagcaggtt
5521 ttcatgtact acaaccattt tcatggtcct atccctgcat cgattggtaa tgcttccagc
5581 atctcaatat ttaccattgg tttaaactct tttagcggtg ttgttccacc ggagattgga
5641 aggatgagaa atcttcagag actagagctt ccagaaactc ttttggaagc tgaagaaaca
5701 aatgattgga aattcatgac ggcattgaca aattgctcca atcttcaaga agtggaactg
5761 gcaggttgca aatttggtgg agtcctccct gattctgttt ccaatctttc ctcttcgctt
5821 gtatctctct ccattagaga taacaaaatt tcagggagct tacctagaga tatcggtaat
5881 ctcgttaatt tacaatatct ttctctcgct aacaactcct tgacaggatc ccttccctct
5941 tccttcagca agcttaaaaa tttacgtcgt ctcactgtag ataacaacag gttaattggt
6001 tctctcccat tgactatcgg taatcttaca caactaacta atatggaggt ccaatttaat
6061 gcctttggtg gtacaatacc aagcacactt ggaaacctga ccaagctgtt tcaaataaat
6121 cttggccaca ataactttat agggcaaatt cccattgaaa tatttagcat tcccgcactc
6181 tctgaaattt tggatgtgtc ccataataac ttggagggat caataccaaa agaaataggg
6241 aaacttaaaa atattgtcga attccatgct gattcgaaca aattatcggg tgagatccct
6301 agcaccattg gtgaatgcca acttctgcag catcttttcc tgcaaaacaa tttcttaaat
6361 ggtagcatcc caatagctct gactcagttg aaaggtctgg acacacttga tctctcaggc
6421 aacaatttgt caggtcagat acctatgtcc ttaggggaca tgactctgct ccactcgctg
6481 aacctttcgt tcaacagctt ccacggtgaa gtgccaacca atggtgtttt tgcaaatgct
6541 tctgaaattt acatccaagg caatgcccat atttgcggtg gcatacctga actacatctt
6601 ccgacgtgtt ccttaaaatc aagaaagaaa aggaaacatc aaattctgct gttagtggtt
6661 gttatctgtc tcgtttcgac acttgccgtc ttttcgttac tctacatgct tctaacctgt
6721 cataagagaa gaaagaaaga agtccctgca acgacatcca tgcaaggcca cccaatgatc
6781 acttacaagc agctggtaaa agcaacggat ggtttttcgt ccagccattt gttgggttct
6841 ggatcttttg gctctgttta caaaggagaa tttgatagtc aagatggtga aatcacaagt
6901 cttgttgccg tgaaggtact aaagctagaa actcctaagg cactcaagag tttcacggcc
6961 gaatgcgaaa cactacgaaa tacgcgacac cggaatcttg tcaagatagt tacgatttgc
7021 tcgagcatcg ataacagagg gaatgatttc aaagcaattg tgtatgactt catgcccaat
7081 ggcagtctgg aagattggct acaccctgaa acaaatgatc aagcagagca aaggcacttg
7141 actctgcatc agagagtgac catactactt gatgttgcat gtgcattgga gcatcttcac
7201 ttccatggcc ctgaacctat tgtacactgt gatattaaat caagcaatgt gttgttagat
7261 gctgatatgg tagctcatgt tggagacttt ggacttgcaa gaatacttgt tgagggaagc
7321 tcattgatgc aacagtcaac aagttcgatg ggaatcaggg ggacaattgg ttacgcagca
7381 ccaggttaat cctaaactgt ttatgtctac ctcctttcat tgtttttttt ttagatttgc
7441 tctggtccaa caaaaaatac ctaaagatac agatacttgt acctcacagt actaaatagt
7501 ttttgatcat tgcattgtta gatccaacga tcagaaaacg atttggtacc gtgaccgtga
7561 ggtatcggaa tctcgagata tttttttgtt cgaccgtagc aaatctattt ttttgtttgt
7621 tttcttctct ttaatgtttt atgactatga aataattttt atttctggaa aacagagtat
7681 ggtgtcggga acactgcctc gacacatgga gatatttaca gttatggaat tctagtgttg
7741 gaaacagtaa ccgggatgcg gccggcagac agtacattca gaactggatt gagcctccgt
7801 cagtacgttg aaccgggtct acatggtaga ctaatggatg ttgttgacag gaagcttggt
7861 ttggattccg agaaatggct tcaggctcga gatgtttcgc cacgcagcag tattactgaa
7921 tgccttgttt cactgcttag acttgggctg tcttgctctc aggaattgcc atcgagtaga
7981 acgcaagccg gagatgtcat caatgaactg cgtgccatca aagagtctct ctcgatgtca
8041 tccgacatgt gaagatgtga gacatgctga tgttatgttg gagtatttcg ttgtaatgta
8101 atgtgaaggg tgagtgtgtg actgcttggt tgtaagctat ttcctgatct gcccatcaga
8161 tcatgtatct gttctattgt tgtatttctc agaacaacca cacacctaag taggagtaca
8221 caatagtgta tttgtgtgat ttcaatattg gtgcataccc atgctatgtg aacagtcaat
8281 cggggagcga ttcacaccat accgtgaaat cgacctaatc agctaatcta attctacagg
8341 ctgcctttgc atgacagtgt gatattaaat tagcccagcc ctttttagca aacgatggga
8401 gggtcaatgc tctaga
```

SEQ ID NO:5

///translation="MISLPLLLFVLLFSALLLCPSSSDDDGDAAGDELALLSFKSSLL

YQGGQSLASWNTSGHGQHCTWVGVVCGRRHPHRVVKLRLRSSNLTGIISPSLGNLSFL

RTLQLSNNHLSGKIPQELSRLSRLQQLVLNFNSLSGEIPAALGNLTSLSVLELTNNTL

SGSIPSSLGKLTGLYNLALAENMLSGSIPTSFGQLRRLSFLSLAFNHLSGAIPDPIWN

ISSLTIFEVVSNNLTGTLPANAFSNLPNLQQVFMYYNHFHGPIPASIGNASSISIFTI

GLNSFSGVVPPEIGRMRNLQRLELPETLLEAEETNDWKFMTALTNCSNLQEVELAGCK

FGGVLPDSVSNLSSSLVSLSIRDNKISGSLPRDIGNLVNLQYLSLANNSLTGSLPSSF

SKLKNLRRLTVDNNRLIGSLPLTIGNLTQLTNMEVQFNAFGGTIPSTLGNLTKLFQIN

LGHNNFIGQIPIEIFSIPALSEILDVSHNNLEGSIPKEIGKLKNIVEFHADSNKLSGE

IPSTIGECQLLQHLFLQNNFLNGSIPIALTQLKGLDTLDLSGNNLSGQIPMSLGDMTL

LHSLNLSFNSFHGEVPTNGVFANASEIYIQGNAHICGGIPELHLPTCSLKSRKKRKHQ

ILLLVVVICLVSTLAVFSLLYMLLTCHKRRKKEVPATTSMQGHPMITYKQLVKATDGF

SSSHLLGSGSFGSVYKGEFDSQDGEITSLVAVKVLKLETPKALKSFTAECETLRNTRH

RNLVKIVTICSSIDNRGNDFKAIVYDFMPNGSLEDWLHPETNDQAEQRHLTLHQRVTI

LLDVACALEHLHFHGPEPIVHCDIKSSNVLLDADMVAHVGDFGLARILVEGSSLMQQS

TSSMGIRGTIGYAAPEYGVGNTASTHGDIYSYGILVLETVTGMRPADSTFRTGLSLRQ

YVEPGLHGRLMDVVDRKLGLDSEKWLQARDVSPRSSITECLVSLLRLGLSCSQELPSS
RTQAGDVINELRAIKESLSMSSDM"

SEQ ID NO:6

DEFINITION Oryza longistaminata receptor kinase-like protein (Xa21) gene, complete cds and family member C, pseudogene.
ACCESSION U72723
SOURCE long-staminate rice.
ORGANISM Oryza longistaminata
Eukaryotae; mitochondrial eukaryotes; Viridiplantae; Charophyta/Embryophyta group; Embryophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.

FEATURES	Location/Qualifiers
source	1..19639
	/organism="Oryza longistaminata"
	/strain="IRBB21"
	/chromosome="11"
	/map="11q, RG103"
gene	5213..18201
	/gene="Xa21"
CDS	join(5213..7889,8732..9132)
	/gene="Xa21"
	/note="disease resistance gene"
	/codon_start=1
	/product="receptor kinase-like protein"
3' flanking	9132..15118
misc_feature	9645..9769

```
                     /note="Pop-O11, transposon-like element"
                     13040..13248
                     /note="Ds-rice1, transposon-like element"
          CDS        join(15118..17720,17827..18201)
                     /gene="Xa21"
                     /note="family member C; 2 bp deletion causing a
                     frame-shift of the ORF compared to family member A1"
                     /codon_start=1

          misc_feature  16183..16184
                     /gene="Xa21"
                     /note="location of a 2 bp delection compared to family
                     member A1"

                     BASE COUNT     5380 a    4394 c    3800 g    6065 t
ORIGIN
        1 aagctttgct catttcttct ccaattacaa ttaatcgtgt gtgcatcaag taattaatta
       61 agaaactctg gcttgttgaa aggtcgcagt gacaattaat cgtgagtgca taggatgggg
      121 aaaacacgag ggatcggtcg accatcgggg ggagcaaaaa tcaagcgctc ccccgccccca
      181 cgacgcgcac atgccacgcc accccaccac gcaccacgtg tgccgttgta aacacctgcc
      241 acgtgtcccc atcacatact ccgttgtaac aaatcaccca tatattttgg aaaccctata
      301 ttaggagaat tcgtttcatt ttttttctac caaaaatatt tcacctagtg tactcacaat
      361 gtttcactat gtatagatct aattttgtag taaattgaaa cattcttttg caataaatta
      421 cccatatatt atggaaaccc tctattaagg gaattcgttt catttttat tccactaaaa
      481 atgtttcgcc tagtgtactt gtaatgtttc actatgtatg gattaaatgt tgcagtgact
      541 tgaaacattc tttcgctatt tgctgaaaca ttgtttttat ataaggtgaa acagcgcccg
      601 atttaaacga ttgaaatatt ttcgatctac ttagtgaaac aattccaata tacttggtgc
      661 aacaacgtgc aacatttaat tataattcca taataagctt gtcgacattt gatagtcacg
      721 actagggtat ttggatggta tggggatcat cggtaccagg gtatatgcga gattgaggta
      781 aaagagatgg agatagggat ttttatatag gttcgggccc cttatctgat aggtaatagc
      841 cctacatcct gtttatatgt gattgatata gaaaaaccac agaatacaac aattgggata
      901 acctatctag ccgttgttga cttggcggca cgaacaccaa ctcgtagtcg acgacgggga
      961 agcctttctc ctcgattgtg aactcgacaa gattagagat atcgctagat ccctcttgcc
     1021 ggcctctgta ggtaccggat ggggtgtgtc taggctaatc tcagatgtcg atgtttggcg
     1081 gcgtattggc ttgtgtcttg tggcttctat gttgtgtgtc ccctctcctc caatggaggc
     1141 ttggatttag actcatagat ttccccttgt ccaagtagaa ctagggagac caatatagat
     1201 acaatccgag tagtacttgt cgtttctata tagaactcta ttttgtcctt ccttatccgg
     1261 aactccttct atatatgagg tatgtttccg tataagactt ggtatgtggt gggcctcgcc
     1321 gagcttagtc gattactatt gggtatgtgg tatcctaggc cccagctgcc attttccaca
     1381 aagacagctt gaatcttata acaaaccttg ctgaagctga caaatcctag cccccagcca
     1441 tgaagttgga aaaatcaatt tccgattaca caaattggtt aatacgcaac catttagtgc
     1501 tcttaacatg accaggtttt acatgttcat tcggctctta gaatctgaca agaccttatc
     1561 tgctctgggc gtccccagcc gaaattccat tagttttctc ggaggcttgt cagaacagcg
     1621 taaagggaca ataggactgc cttcaagatg aggcgatata agaggggatc aacagacaaa
     1681 tattgcacat ataaaactta cagaagttga tgtagatgat gagacgacca ccacactagg
     1741 caaagaccag gtgtatagtt gtactcaaca aatcgcagag gtagtgagag atcgctacga
     1801 tctactggtc gaagatcagg tgtaggcgta ttctcgatca cctgaagaag aatctttagg
     1861 tgttgagaga tcgctactat ctactggtca acactagtaa aaaaacctca tagagatcgg
     1921 cactataggt gccgaacagc taaaacccgc acctatatac tttcctctcg tggactcaaa
     1981 gcacgtaaac cgacacttta gcaactatag gtgcggtcta aagagaccga cactatagta
     2041 tagtgctggt tttaaaaaac ccgacacttt aatataatat agtgtcggtc tctttaaaac
     2101 gacacaatat aaatatacgt gtcggttttt aataaacggc acctatcaaa ccgagcctag
     2161 ctgtcgagtc gagccaatcc aggtgacgca tatattagtc tcgtccttat cgcctccggt
     2221 ctctctctct ctctctctct ctctttctct ctcgcctctc tgtgtagcgc gcggcggtgg
     2281 tcgggcagcg tcccggcgta ggcaacagtg gtggcgagat gggaggcggt ggcgcatcac
     2341 agctattcac tttagcgcct tgaataataa tcggcatcat gatctactta tgcttcgtgc
     2401 aagagggaag aatgtagata cacatggtca cagctaagtg ctatgatcgc ggctcatctc
```

```
2461 tccaaataga ttcatgccat ccgtacttaa acagaacctt atgttccgtg catccttttct
2521 gaagtcttga aatgttttat cgatgtttta ccactacgac ccatctgtag ggtgtctcgc
2581 atcccgtctt gtttatgttt ttatgtgtgc cagcgcacca ttctagcacg cgctttgttc
2641 ctgaacaaac accttagccg tggtattaaa tggaaatact gcatgatcat agcaggaact
2701 ctcttcttcg aaggatgtcc gtcaacatca cctaggtcat ctcgtctaat cttatatcgt
2761 agtgccttac aaaccaggcg tgcttctagg ttctcgtact cctccacgcg atagaggata
2821 caatcattca gacatacgtg tatcttctgt acttccagtc cgagagggca gattaccttt
2881 ttagcttcgt acgttgtttc gagcaattcg tttccctcgg gaagaatatt ctttacgagt
2941 ttcaataact cgccaaatgc cttgtcagtc acaccatatt ttgccttcca ttgtcagaat
3001 tccagtgtgg tacccaactt tttgtgctcc tattcgcaac ctgggtacaa ggacgttttg
3061 tggtatgcta acatccggtc caatttctag accaactttt tcacttttgc agtcattttt
3121 tacatcgcac aacatttgac caagattatc cgtaccgccg ttttcttcaa cagctctgtc
3181 cgcctcgcct attgtatttt ctgcaaatcc atcatattga gcccagtctg aaatgttgtc
3241 gtcttcttat taatttttctt ctattgcaac ccctaactct ccgtgcaagg tctaacaatt
3301 atagccaggt atgaatcccg attccaacaa gtggatatga agacttctta gatgtagaat
3361 actccttctg gtttttgcac ttttgcatgg acaacatata aaacccttag gcttgtgggc
3421 attggccaca ctcaaaaaat aatgcacgtc atcaataaac tccttcgacc gtcgttctgc
3481 agtgtacatc cattaccgat tcatctacat taggagataa taattgtaaa ggaagtccac
3541 aaaagtgaaa ctacttaaat aatcatataa taatttaaaa tatcataatt aaattgaaaa
3601 ctgacggttt taatgtattc ttcttatttc taacgatttt aacgagtttt aaatggactt
3661 aatcggagtc atgattaact atttataaat tttatctgtc tcaataaatt gtaaatatat
3721 ttttccatgt attatccttg tttaattatt tttaaaagtt ctaaacatat ttttaatgca
3781 ttctacttat ttctaactat tttaaagatt ttcaaatgga cttagttttc tatttttatt
3841 cttcatttta tatatttgcc ctttgttgtc tctttttaac aattttatac aaatatttat
3901 aattttatta agtaccctaa ttttccctaa acaatttttc tctctcatcg tatttccata
3961 tatcttttta agataataat ggatataaac atagctagaa atgtaaatgt tcaccttgca
4021 tcaatagggg atgaagttgc taacctttta gatctcctcg atttgtataa tataaccaaa
4081 atattttcac caaaaatttc gttaaacatc cgagatattt gttgtttttg ccgatcgagc
4141 aaagattagt agtccagcag tgtctgcacc accaccatcg tgataatgca tcttgtgtgt
4201 tattcttgat gagaaaatac gtagtgaaaa ccacatatgt ggtggaaact tagaaactac
4261 cgttagatcg agaaatggat gtccaagatt cgtccacgtc accaagagat aaaatttagc
4321 tcgcagattc acttatgagt taaaatttta atgagagtta aattttaact catgttgatg
4381 tggacgaata tcggacatcc atttctcgat ccaacgatag cttccaagtt tccactacat
4441 atgtggtttg cactatatat tttcccattc ttgattatgt gtttgagagc agctagcaca
4501 aagagaaaaa aaagcatcgt ttttcacgcg tatgttttca gaactgttaa atggtgtgtt
4561 ttttgaaaaa actttctata gaaaagtttc tttaaaaaat atattaatct attttttaag
4621 tttaaaataa ttactactta attaattata cactaacagc ttatttcgtt ctacgtatct
4681 tgtcaatttt cgctattcct ttcttctcaa acacggcatt ggatgctctc atagcacttg
4741 ctcgttcgga tagaagactt gacgaagacg accgctacaa cttggtgtgt tatatcgtgc
4801 tttgtttagc ataatcatta catatattcc atgccgaagt gccgacgatg agaccgtgtt
4861 cgatgcatct ttgtatggca tctagggaca aagagcatag agtccctacc atagtaccag
4921 ctcgcgcaga agacttgacg agaagaccga ctgctacacc ttggtgtgta ataatatcgt
4981 gttgtgtgta ccatgcatac tcctttaaaa caaataatgg tggtaacagt aaatctgtca
5041 tcccacccac tctcattgta aattttgcaa gttctcactt gaacttctta atactccatc
5101 cgtttgcgtg tgttctttca gaatttgcgt gagcactttt tcttctatat aatctgtcta
5161 gtccatgagc taaaccaaca tctctcgctg tcttgccttg cacttctgca cgatgatatc
5221 actcccatta ttgctcttcg tcctgttgtt ctctgcgctg ctgctctgcc cttcaagcag
5281 tgacgacgat ggtgatgctg ccggcgacga actcgcgctg ctctctttca agtcatccct
5341 gctataccag gggggccagt cgctggcatc ttggaacacg tccggccacg gccagcactg
5401 cacatgggtg ggtgttgtgt gcggccgccg ccgccgccgg cacccacaca gggtggtgaa
5461 gctgctgctg cgctcctcca acctgtccgg gatcatctcg ccgtcgctcg gcaacctgtc
5521 cttcctcagg gagctggacc tcggcgacaa ctacctctcc ggcgagatac caccggagct
5581 cagccgtctc agcaggcttc agctgctgga gctgagcgat aactccatcc aagggagcat
5641 ccccgcggcc attggagcat gcaccaagtt gacatcgcta gacctcagcc acaaccaact
5701 gcgaggtatg atcccacgtg agattggtgc cagcttgaaa catctctcga atttgtacct
5761 ttacaaaaat ggtttgtcag gagagattcc atccgctttg ggcaatctca ctagcctcca
5821 ggagtttgat ttgagcttca acagattatc aggagctata ccttcatcac tggggcagct
5881 cagcagtcta ttgactatga atttgggaca gaacaatcta agtgggatga tccccaattc
```

```
5941 tatctggaac ctttcgtctc taagagcgtt tagtgtcaga gaaaacaagc taggtggtat
6001 gatccctaca aatgcattca aaacccttca cctcctcgag gtgatagata tgggcactaa
6061 ccgtttccat ggcaaaatcc ctgcctcagt tgctaatgct tctcatttga cagtgattca
6121 gatttatggc aacttgttca gtggaattat cacctcgggg tttggaaggt taagaaatct
6181 cacagaactg tatctctgga gaaatttgtt tcaaactaga gaacaagatg attgggggtt
6241 catttctgac ctaacaaatt gctccaaatt acaaacattg aacttgggag aaaataacct
6301 gggggggagtt cttcctaatt cgttttccaa tctttccact tcgcttagtt ttcttgcact
6361 tgaattgaat aagatcacag gaagcattcc gaaggatatt ggcaatctta ttggcttaca
6421 acatctctat ctctgcaaca acaatttcag agggtctctt ccatcatcgt tgggcaggct
6481 taaaaactta ggcattctac tcgcctacga aaacaacttg agcggttcga tcccgttggc
6541 cataggaaat cttactgaac ttaatatctt actgctcggc accaacaaat tcagtggttg
6601 gataccatac acactctcaa acctcacaaa cttgttgtca ttaggccttt caactaataa
6661 ccttagtggt ccaataccca gtgaattatt caatattcaa acactatcaa taatgatcaa
6721 tgtatcaaaa aataacttgg agggatcaat accacaagaa atagggcatc tcaaaaatct
6781 agtagaattt catgcagaat cgaatagatt atcaggtaaa atccctaaca cgcttggtga
6841 ttgccagctc ttacggtatc tttatctgca aaataatttg ttatctggta gcatcccatc
6901 agccttgggt cagctgaaag gtctcgaaac tcttgatctc tcaagcaaca atttgtcagg
6961 ccagataccc acatccttag cagatattac tatgcttcat tccttgaacc tttctttcaa
7021 cagctttgtg ggggaagtgc caaccattgg tgctttcgca gctgcatccg ggatctcaat
7081 ccaaggcaat gccaaactct gtggtggaat acctgatcta catctgcctc gatgttgtcc
7141 attactagag aacagaaaac atttcccagt tctacctatt tctgtttctc tggccgcagc
7201 actggccatc ctctcatcac tctacttgct tataacctgg cacaagagaa ctaaaaaggg
7261 agccccttca agaacttcca tgaaaggcca cccattggtc tcttattcgc agttggtaaa
7321 agcaacagat ggtttcgcgc cgaccaattt gttgggttct ggatcatttg gctcagtata
7381 caaaggaaag cttaatatcc aagatcatgt tgcagtgaag gtactaaagc ttgaaaatcc
7441 taaggcgctc aagagtttca ctgccgaatg tgaagcacta cgaaatatgc gacatcgaaa
7501 tcttgtcaag atagttacaa tttgctcgag cattgataac agagggaacg atttcaaagc
7561 aattgtgtat gacttcatgc ccaacggcag tctggaagat tggatacacc ctgaaacaaa
7621 tgatcaagca gaccagaggc acttgaatct gcatcgaaga gtgaccatac tacttgatgt
7681 tgcctgcgca ctggactatc ttcaccgcca tggccctgaa cctgttgtac actgtgatat
7741 taaatcaagc aatgtgctgt tagattctga tatggtagcc catgttggag attttgggct
7801 tgcaagaata cttgttgatg ggacctcatt gatacaacag tcaacaagct cgatgggatt
7861 tatagggaca attggctatg cagcaccagg tcagcaagtc cttccagtat tttgcatttt
7921 ctgatctcta gtgctatatg aaatagtttt tacctctagt gaaactgatg gagaatataa
7981 gtaattaatt gaactaatta aattgcacaa aaataagatt atttgccata tctattcaga
8041 tgctaaatat agctagttca tagaggtaca gatttttta tataggactc tagagctacc
8101 acacactcaa atcaaattat gggtgttttc tgctctacac tgcaatatga aatgattatt
8161 acttctacat gaactgatgg aggagtttca gaaggatcaa atttgagtaa atttttcaat
8221 tctacattta agaaacactt ttttttcata tgctagttac attttttat ttcacgagct
8281 tacattgacc atgaaaaata cttggcacta cttactaatt cccacatgga ggtagtgaaa
8341 ataatataga tacaaaaacg aaatatccta tgttgtgtga tatactataa tcacaatgaa
8401 cacaaacagg attcgtacaa aagtaattag ccatcatagc aactgattgc ttggggtaac
8461 tgtatagcac aatcatacca aatttcttta gatatgtatc tgtaaattag attcttaaag
8521 ttaaatatga aatttcattg gtatttatgt ttctttatat aataaaaatt aatccagcct
8581 ttgcatctat catttgtcca gacatccttg ttatttgtga tatttaacac gtaaatttac
8641 ataattatac atccaagttc tttttattta acactgtaaa tttcaaatcg tacatgttat
8701 aaagaatgta ctatatttcc tgctcaaaca gagtatggcg ttgggctcat tgcatcaacg
8761 catggagata tttacagcta tggaattcta gtgctggaaa tagtaaccgg gaagcggcca
8821 actgacagta cattcagacc cgatttgggc ctccgtcagt acgttgaact gggcctacat
8881 ggcagagtga cggatgttgt tgacacgaag ctcattttgg attctgagaa ctggctgaac
8941 agtacaaata attctccatg tagaagaatc actgaatgca ttgtttggct gcttagactt
9001 gggttgtctt gctctcagga attgccatcg agtagaacgc caaccggaga tatcatcgac
9061 gaactgaatg ccatcaaaca gaatctctcc ggattgtttc cagtgtgtga aggtgggagc
9121 cttgaattct gatgttatgt ctcgtaatgt tttattgcca cacttcagat cgacttctgc
9181 agtggtatct accacacgat cactaaagtc accgtggcta tttcctgatc cagcatatct
9241 gatcatgcat gttctgtgtt gtatacctgt attttactct gaattgccac accgcaaccc
9301 tgcctctgtt tgtttggtat acaaaagata gtgatgagtt tattgtttta ggggcttcct
9361 agttggcgcg tgtgcatgcc ggcatgcacg cagcccgagg gtgggtttct ttttttcca
```

```
 9421 ttgttattcc gttgcttttt ttcaccacgg tagatttttt tttccggatt tccatttttt
 9481 ccgttgtttt tctctatcgc ttatgttggc ggatttttt ccgtggtttt ctttccgaag
 9541 acgagtatat ctaacgtaac taacatgtta cttttagata acgatggtta ttaagataag
 9601 atttttctct ggaagatttt tgtaagtaac agattgaaaa caaatctata cgtgaggtca
 9661 aattttgaaa actttcaatc tagatttaaa agcttttcaa ctcaaaattt gaatttttga
 9721 agtgaaaatt tgaatacttt caaaaattac tagtaatcga caaaaaaaat atggaaatgg
 9781 aaacggaaat agttttgctg ttataccgat cgtttccata tttaccgtat tcttatagaa
 9841 attaccgttt cttataatat ggtaattacc gtatttctaa atatgttgat atttataggg
 9901 catgtctcta cttgactcac agtttagaga ttgattgact atttaatcaa atccctaact
 9961 tgattgcatg gctaaaatgg agttgatttc taatttatat agtatagctt gaatttattt
10021 gtaaatataa catacttatg taaagttaaa tatatgtttt ctatagttta atgtttctgt
10081 atttgttacc ggttttcgat ctgtaccgac atatttccat cagtattatt ccatttccgg
10141 ttttccgata tttccgatat cgttttcgtt tccgacttta ccgttttcga tttcatttcc
10201 gagaaaaata tgattatgga aatggtcgag gctgttttcc gatcgtttcc gaccgttttc
10261 atccctaccc gtagtaataa tatataacat tttatctcta atctttctct ctctcatatc
10321 aatgaataat cgctaagaga ctgctattaa caaggcttat atatatatat gccgtcgatc
10381 agtcattttg aaacggccca cttcttttcc atctatatgc attcatgaaa tacatggtat
10441 atcccatcga tcggacatca cctgttagcg cgtacgccat cgtcgtcatc aacctagcta
10501 gggcaaacgc accttgctga gctccgatcc tccgatcgcc accatcacca atgaacaagc
10561 tgctgcggcc tctcggtggc ctgaggttgc tcaaccgaga agaacatccg ttccgatgct
10621 tctcctcctc catcgatctc gtcttcccag gtcgccgccg ccgccacatg gcaaccaccg
10681 tgacccaccc gccgccgacg gaatccgctg gttcgacggc ggcggccgcg actgctgacc
10741 cggcctcggt gatgctggaa cattggggct gcctcagggg ctccacgccg gcgaacgtag
10801 tcgccgacga caacaccgcc gcggagtccc gcacctcccg cggccaaccc ctccgggtcg
10861 ccctcgcccg cgcgtcgccg ccggcgatct ccttcatctg cttcgatcgc ggggatgatg
10921 gctacgtcat cgcggctcac ggcgactctg tcctcttccg gatgagttgg aacgactact
10981 tcgtctacat ggccgccggc ggccgccgtc gctgacgctg ctccccgtct gcgacatccc
11041 catgaacgag cgctgctggg tcagcaagga ccgtttcaag gacagcttcc acaccacggg
11101 ccgggagttc gaccagcagg acaccggcat cctgcgcctc cgcggcgacg acggcggcga
11161 ggaggcgccg ctctagtggc gcagtccaga tcgcgcacga gccgccgttc gacacggccg
11221 agctctgcgt gctccgcccc ggccacggcg agtgggagct caagatggcg gtgcccatcg
11281 tccaccatga cagcttgaat cttataacaa accttgctga agctgacaaa tcctagcccc
11341 cagccatgaa gttggaaaaa tcaatttccg attacacaaa ttggttaata cgcaaccatt
11401 tagtgctctt aacatgacca ggttttacat gttcgttcgg ctcttagaat ctgacaatac
11461 cttatctgct ctgggcgtcc ccagccgaaa ttccattagt tttctcggag gcttgtcaga
11521 acagcgtaaa gggacaatag gactgccttc aagatgaggc gatataagac gggatcaaca
11581 gacaaatatt gcacatataa atacttacag aagttgatgt agatgatgag acgaccacca
11641 cactaggcaa agaccaggtg tatagttgta ctcaacaaat cgcagaggta gtgagagatc
11701 gctacgatct actggtcaaa gatcaggtgt aggcgtattc tcgatcacct gaagaagaat
11761 ctctaggtgt tgagagatcg ctactatcta ctggtcaaaa ctagtaaaaa aacctcatag
11821 agatcggcac tataggtgcc gaacagctaa aacccgcacc tatatacttt cctctcgtgg
11881 actcaaagca cgtaaaccga cactttagca actataggtg cggtctaaag agaccgacac
11941 tatagtatag tgctggtttt aaaaaacctg acactttaat ataatatagt gtcggtctct
12001 ttaaaacgac acaatataaa tatacgtgtc ggttttaata aacggcacct atcaaacgat
12061 cctagctgtc gagtcgagcc aatccaggtg acgcatatat tagtctcgtc cttatcgcct
12121 catctctctc tctctctctc tctctctctt tctctctcgc ctctctgtgt agcgcgcggc
12181 ggtggtcggg cggcatcccg gcgtaggcaa cagtggtggc gagatgggag gcggtggcgc
12241 atcacagcta ttcactttag cgccttgaat aataatcggc atcatgatct acttatgctt
12301 cgtgcaaggg ggaagattgt agatacacat ggtcacagct aagtgctatg atcgcggctc
12361 atctctccaa atagattcat gccatccgta cttaaacaga accttatgtt ccgtgcatcc
12421 tttctgaagt cttgaaatgt tttatcgatg ttttaccact acgacccatc tgtagggtgt
12481 ctcgcatccc gtcttgttta tgtttttatg tgtgccagcg caccattcta gcacgcgctt
12541 tgttcctgaa caaacacctt agccgtggta ttaaatggaa atactgcatg atcatagcag
12601 gaactctctt cttcgaagga tgtccgtcaa catcacctag gtcatctcgt ctaatcttat
12661 atcgtagtgc cttacaaacc aggcgtgctt ctaggttctc gtactcctca acgcgataga
12721 ggatacaatc attcagacat acgtgtatct tctgtacttc cagtctgaga gggcagatta
12781 ccattttagc ttcgtacgtt gtttcgagca attcgtttcc ctcgggaaga atattcttta
12841 cgagtttcaa taactcgcca aatgccttgt cagtcacacc atattttgcc ttccattgtc
```

12901 agaattccag tgtggtaccc aacttttgt gctcctattc gcaacctggg tacaaggacg
12961 ttttgtggta tgctaacatc cggtccaatt tctagaccaa ctttttcact tttgcagtca
13021 ttttttacat cgcacaacat ttgaccaaga ttatccgtac cgccgttttc ttcaacagct
13081 ctgtccgcct cgcctattgt attttctgca aatccatcat attgagccca gtctgaaatg
13141 ttgtcgtctt cttattcatt ttcttctatt gcaacccctа actctccgtg caaggtctaa
13201 caattatagc caggtatgaa tcccgattcc aacaagtgga tatgaagact tcttagatgt
13261 agaatactcc ttctggtttt tgcacttttg catggacaac atataaaacc cttaggcttg
13321 tgggcattgg ccacactcaa aaaataatgc acgtcatcaa taaactcctt cgaccgtcgt
13381 tctgcagtgt acatccattg ccgattcatc tacattagga gataataatt gtaaaggaag
13441 tccacaaaag tgaaactact taaataatca tataataatt taaaatatca taattaaatt
13501 gaaaactgac ggttttaatg tattcttctt atttctaacg attttaacga gttttaaatg
13561 gacttaatcg gagtcatgat taactattta taaattttat ctgtctcaat aaattgtaaa
13621 tatatttctc catgtattat ccttgtttag ttatttttaa aagttctaaa catattttta
13681 atgcattata cttatttcta actattttaa agattttcaa atggacttag ttttctattt
13741 ttattcttca ttttgtatat ttgcccttg ttgtctcttt ttaacaattt tatacaaata
13801 tttataattt tattaagtac cctcattttc cctaaacaat ttttctctct catcgtattt
13861 ccatatatct ttttgagata ataatggata taaacatagc tagaaatgta aatgttcacc
13921 ttgcatcaat aggggatgaa gttgctaacc ttttagatct cctcgatttg tataatataa
13981 ccaaaatatt ttcaccaaaa atttcgttaa acatccgaga tatttgttgt ttttgccgat
14041 cgagcaaaga ttagtagtcc agcagtgtct gcaccaccac catcgtgata atgcatcttg
14101 tgtgttattc ttgatgagaa aatacgtagt gaaaaccaca tatgtggtgg aaacttggaa
14161 actaccgtta gatcgagaaa tggatgtcca agattcgtcc acatcaccaa gagataaaat
14221 ttaactcgca gattcactta tgagttaaaa ttttaatgag agttaaattt taactcatgt
14281 tgatgtggac gaatatcgga catccatttc tcgatccaac gatagcttcc aagtttccac
14341 tacatatgtg gtttgcacta tatattttcc cattcttgat tatgtgtttg agagcagcta
14401 gcacaaagag aaaaaaaagc atcgtttttc acgcgtatgt tttcagaact gttagatggt
14461 gtgtttttg aaaaaacttt ctatagaaaa gtttctttaa aaaatatatt aatctatttt
14521 ttaagtttaa aataattact acttaattaa ttatacacta acagcttatt tcgttctacg
14581 tatcttgtca attttcgctc atcctttctt ctcaaacacg gcattggatg ctctcatagc
14641 acttgctcgt tcggatagaa gacttgacga agacgaccgc tacaacttgg tgtgttatat
14701 cgtgctttgt ttagcataat cattacatat attccatgcc gaagtgccga cgaggagacc
14761 gtgttcgatg catctttgta tggcatctag ggacaaagag catagagtcc ctaccatagt
14821 acctgctcgc gcagaagact tgacgagaag accgactgct acaccttggt gtgtaataat
14881 atcgtgttgt gtgtaccatg catactcctt taaaacaaat aatggtggta acagtaaatc
14941 tgtcatccca cccactctca ttgtaaattt tgcaagttat cacttgaact tcttaatact
15001 ccatccgttt gcgtgtgttc tttcagaatt tgcgtgagca ctttttcttc tatataatct
15061 gtctagtcca tgagctaaac caacatctct cgctgtcttg ccttgcactt ctgcacgatg
15121 gtatcactcc cattattgct cttcgtcctg ttgttctctg cgctgctgct ctgcccttca
15181 agcagtgacg acgatggtga tgctgccggc ggcgaactcg cgctgctctc tttcaagtca
15241 tccctgctat accaggggggg ccagtcgctg gcatcttgga acacgtccgg ccacagccaa
15301 cactgcacat gggtgggtgt tgtgtgcggc cgccggcacc cgcacagggt ggtgaagctg
15361 cggctgcgct cgtccaacct gaccgggatc atctcgccgt cgctgggcaa cctatccttc
15421 ctcaggacgc tgcaactcag caacaaccac ctgtccggca agatacccca ggagctcagc
15481 cgtctcagca ggctccagca actggtactg aatttcaaca gcctatcggg tgagattcca
15541 gctgctttgg gcaatctaac cagtctctcg gttcttgtgc tgactaacaa tacactgtct
15601 ggttctatcc cttcatccct gggcaagctc accggcctct ataatcttgc actggctgaa
15661 aatatgctgt ctggttccat cccttcatct ttcggccaat tgcgcagatt atctttcctt
15721 agcttagcct tcaaccactt aagtggagca atcccagatc ctatttggaa catctcctct
15781 ctcaccatat ttgaggtcat atccaacaag ctaaatggta cactgcctac aaatgcattc
15841 agtaatcttc ctagtctgaa ggaggtatac atgtattaca accagtttca tggtcatatc
15901 ccggcatcga taggtaatgc ttccaacatc tcaatattta ccattggttt aaactccttt
15961 agcggtgttg ttccactgga gattggaagg ctgagaaatc ttcagaggct agagcttgga
16021 gaaactcttc tagaatctaa agaaccaaac gattggaaat tcatgatggc attgacgaat
16081 tgctccaatc ttcaagaagt agaattggga ctttgtaaat ttggtggagt cattcctgat
16141 tctgtttcca atctttcctc ttccctatta tatctctttt ttcgataaca taatttcagg
16201 gagcttacct aaggatatcg gtaatctcgt taatttagaa actctttctc tcgctaacaa
16261 ctccttgaca ggatcccttc cctcatcctt cagcaagctt aaaaatttac atcgtctcaa
16321 acttttaac aacaaaataa gtggttctct cccattaacc attggtaatc ttacacaact

```
16381 aactaatatg gagctccact ttaatgcctt cggtggtaca ataccaggca cacttggaaa
16441 cctgaccaag ttgtttcaaa taaatcttgg ccataataac tttataggtc aaattcccat
16501 tgaaatattt agcattcctg cactctctga aattttggat gtgtctcata ataacttgga
16561 gggatcaata ccaaaagaaa tagggaaact taaaaatatt gtcgaattcc atgctgattc
16621 gaacaaatta tcgggtgaga tccctagcac cattggtgaa tgccaacttc tgcagcatct
16681 tttcctgcaa aacaatttct taaatggtag catcccaata gctctgactc agttgaaagg
16741 tctggacaca cttgatctct caggtaagaa tttgtcaggt cagataccta tgtccttagg
16801 ggacatgcct ctgctccact cgctgaacca ttcgttcaac agcttccacg gtgaagtgcc
16861 aaccaatggt gtttttgcaa atgcttctga aatttacatc caaggcaatg cccatatttg
16921 cggtggcata cctgaactac atcttccgac gtgttcctta aaatcaagaa agaaaaagaa
16981 acatcaaatt ctgctgttag tggttgttat ctgtctcgtt tggacacttg ccgtcttttc
17041 gttactctac atgcttctaa ccggccataa gagaagaaag aaagaagtcc ctacaacgac
17101 atccatgcga ggccacccaa tgatcactta caagcagctg gtaaaagcag cagatggttg
17161 ttcgtccagc catttgctgg gctctggatc ctttggctct gttttcaaag gagaatttga
17221 tagccaagat tgtgaaagca caagtcttgt tgccgtgaag gtactaaagc tggaaactcc
17281 taaggcactc aagagtttca tggccgaatg cgaaacactg cgaaatactc gacacgtcaa
17341 gatagttaca atttgctcga gcatcgataa cagagggaat gatttcaaag caattgtgta
17401 tgacttcatg cccaatggca gtctggaaga ttggctacac cctgaaacaa atgatcaagc
17461 agagcaaagg cacttgactc tgcatcagag agtgaccata ctgcttgatg ttgcatgtgc
17521 attggaccat cttcacttcc atggccctga acctattgta cactgtgata ttaaatcaag
17581 caatgtgttg ttagatgctg atatggtagc ccatgttgga gactttggac tcgcaagaat
17641 acttattgag ggaagctcat tgatgcaaca gtcaacaagt tcgatggtaa tcagggggac
17701 aattggttac gcagcaccag gttaagccta aactgtttat gtctacctca tttcatttct
17761 tctttttttgt gttttcttct ctctagtgtt ttatgactat gaaataattt ttgctactgg
17821 aaaacagagt atggtgtcgg gaacactgcc tcgacacatg gagatattta cagttatgga
17881 attctagtgt tggaaacagt aaccgggaag cggccgacag atagtacatt cagaactgga
17941 ttgagcctcc gtcagtacgt tgaaccgggt ctacatggta gactaatgga tgttgttgac
18001 aggaagcttg gtttggattc cgagaaatgg cttcaggctc gagatatttc gccatgcagc
18061 agtattagtg aatgccttgt ttcactgctt agacttgggt tgtcttgctc tcaggaattg
18121 ccatcgagta gaatgcaagc cggagatgtc atcaatgaac tgcgtgccat caaagagtcc
18181 ctctcgatgt catccggcat gtgaagatgt tggagtattt cgttgtaatg tgatgtgtct
18241 attagtaccc ttcacaactg atttcattct gccgtggtat ttagttattt acaagagagt
18301 cactgaaggg tgagtgtgtg actgcttggc tgtagctatt tcctgatctg cccatcagat
18361 catgtatctg ttctattgtt gtatttctca gaataaccac acacctaagt acacaacact
18421 gtatttgtgt gatttcaata ttgatgcata tatacccatg ctatatgcta gaattatata
18481 caaaaatttt gagatgtctg aagttaacaa tcaatcagga agcgattcac accaaaccgc
18541 gaaatcgacc taatcagcta atctaattgt acatgctgcc tttgcatgac agtgcgatat
18601 taaattagcc cagccctttt tagcaaagga ttggagggtt aatgttctag agaaaaggat
18661 gcttgttagg ttctctcttc tctctcggtt ctcttgttag attatggaac caattgattt
18721 cctctcgaac caatcgattt cgccaccgtc gccaccaggt tccaaatcga tattccggcg
18781 agatgccgtt acagcgtttt atagacgcaa ctcacgccta gactttcttc tcggtacaga
18841 acggccaagc ccaagacatt ccacggccca tttaggcccc ttgtatttag tgttgctttt
18901 actaatgcgc ttctcgcttg ccttggcata atctggaact ccgccttcga tctgtagccc
18961 gtcttcttcc ttgtcgtcct tgaggttctc agtcgcagca cccagcaccc taaacagatg
19021 cctttggtta ggaaacatgg gtggcagcaa cctttccatg ttggaatatc tcgaaatctc
19081 ctacacagat ttattttgct aggtgaaata ccttctcacc ttggtaatta acatttcaaa
19141 tttgtgccac cttaatctca gtggcacata tatgtacaca atagatatct cttggttagc
19201 tcatcaacat ttgcttgaat atctcgacat gacttttata aatctaaggg tacgttcagt
19261 catctccgac acgcaaaacg aagcaccatt cgcgcatgat taatcaaata ttagctaaaa
19321 aattataaaa tggattaata tatttttaa aagccacgct cctataatat tttttaaaaa
19381 atacatagtt taacagtttg aaaagcgtgt cgtacggaaa acgggagagg tgaagttggc
19441 aaagtagact ttagaacaca gcctaagtat ggcagttcat tggcctcaag ttttcaacat
19501 gattccatct ttggaggccc tccttgtttc ctgcagctca cttccaggct cagcccagcc
19561 gctgaacaca actaaacttc acaaaacttg tagtgcttca tgtttcaagg aatgatgact
19621 ttgtcaattc aatggtacc
//
```

SEQ ID NO:7

translation="MISLPLLLFVLLFSALLLCPSSSDDDGDAAGDELALLSFKSSLL
YQGGQSLASWNTSGHGQHCTWVGVVCGRRRRRHPHRVVKLLLRSSNLSGIISPSLGNL
SFLRELDLGDNYLSGEIPPELSRLSRLQLLELSDNSIQGSIPAAIGACTKLTSLDLSH
NQLRGMIPREIGASLKHLSNLYLYKNGLSGEIPSALGNLTSLQEFDLSFNRLSGAIPS
SLGQLSSLLTMNLGQNNLSGMIPNSIWNLSSLRAFSVRENKLGGMIPTNAFKTLHLLE
VIDMGTNRFHGKIPASVANASHLTVIQIYGNLFSGIITSGFGRLRNLTELYLWRNLFQ
TREQDDWGFISDLTNCSKLQTLNLGENNLGGVLPNSFSNLSTSLSFLALELNKITGSI
PKDIGNLIGLQHLYLCNNNFRGSLPSSLGRLKNLGILLAYENNLSGSIPLAIGNLTEL
NILLLGTNKFSGWIPYTLSNLTNLLSLGLSTNNLSGPIPSELFNIQTLSIMINVSKNN
LEGSIPQEIGHLKNLVEFHAESNRLSGKIPNTLGDCQLLRYLYLQNNLLSGSIPSALG
QLKGLETLDLSSNNLSGQIPTSLADITMLHSLNLSFNSFVGEVPTIGAFAAASGISIQ
GNAKLCGGIPDLHLPRCCPLLENRKHFPVLPISVSLAAALAILSSLYLLITWHKRTKK
GAPSRTSMKGHPLVSYSQLVKATDGFAPTNLLGSGSFGSVYKGKLNIQDHVAVKVLKL
ENPKALKSFTAECEALRNMRHRNLVKIVTICSSIDNRGNDFKAIVYDFMPNGSLEDWI
HPETNDQADQRHLNLHRRVTILLDVACALDYLHRHGPEPVVHCDIKSSNVLLDSDMVA
HVGDFGLARILVDGTSLIQQSTSSMGFIGTIGYAAPEYGVGLIASTHGDIYSYGILVL
EIVTGKRPTDSTFRPDLGLRQYVELGLHGRVTDVVDTKLILDSENWLNSTNNSPCRRI
TECIVWLLRLGLSCSQELPSSRTPTGDIIDELNAIKQNLSGLFPVCEGGSLEF"

SEQ ID NO:8

DEFINITION Oryza sativa receptor-like protein gene, family member E, complete cds.
ACCESSION U72724
SOURCE rice.
ORGANISM Oryza sativa
Eukaryotae; mitochondrial eukaryotes; Viridiplantae; Charophyta/Embryophyta group; Embryophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.

FEATURES	Location/Qualifiers
source	1..9424
	/organism="Oryza sativa"
	/strain="IRBB21"
	/chromosome="11"
CDS	2819..5260
	/note="Xa21 gene family member E"
	/codon_start=1
	/product="receptor kinase-like protein"

```
     misc_feature    5211..8128
                     /note="'truncator', an insertion sequence with the
                     characteristics of a transposon"
     misc_feature    5484..5665
                     /note="'Snap-Ol2', transposon-like sequence"
     intron          8357..8644

ORIGIN
        1 aagcttcact tttcccctat tttgttaaga tcagttaatg gtgtcaattc agccataaaa
       61 agacatgtat ccctttaca tttgcccaca aaaaaatttg ctgggtcaca ttagatatac
      121 ggacacacat ttgaagtatt aaacgtagac taatacaaag caaattatat aatccgcatg
      181 taaactgcga gacgaattta ttaagcctaa ttaatccgtc attagcaaat gtttactgta
      241 gcactacatt gttagatcat ggtgcaatta ggcttaaaag atttgtctcg taatttacac
      301 acaaactgtg taattagttt tttcaatatt taatacttcg tacatgtatt taaacgtttg
      361 atgtgacggg acaaaaaaaa aattgccggt gggtctaaat ccaccacccc ccctcctctt
      421 agagttgata tacattttca cacaacttgc acggctgcaa aacttgacta aaaagattca
      481 tgacaaattt ctcgcaaata ttgagatagg aaaagagaga gagaaaatca ggtagcaacg
      541 gcatctcgct ggtctttgtg aacgggaaac tattttaaca gctcgagtgg acgtcaaccc
      601 gtttattgca tgtcatctaa atggttataa aaaaaattga aaaaatatga ataaggatag
      661 atcaatatgt aatatatcaa tccacaaaca tgcaagttaa aatttaactt ctacaagttg
      721 taataaaaat aagaaacaaa actcaaatta ctatatgtat atttacaatt aaatttgtta
      781 tttttgttat aacctgtaga agttaaattt gaacttgcat gtttgtggag tgatatatta
      841 catattgatc tatcttgtca attttttga aaatttttcg taaccatcta gttgacatgc
      901 aataaacgga tggacgtcca ctcgagtgct gctagggcgg atctatagta cgttgtcagg
      961 cgacaccgat gactcccggt aaaaacctat agcaaaactg cttattcata tgttataaca
     1021 ccatgatcta aacataatta gtatactatg ataccgttag acacgttatg acaccaatga
     1081 atcgattttc tggatccgcc acagtgcacg ggattgggat gggaacggcc gatggacgct
     1141 gtgtgttggg ctagctaggt ggggttggga gttgggacta ggtaggttat ttttgccagg
     1201 ggtacaatag ttcttttgca ttttgtaact cttctttttc tttcgagata atccattata
     1261 tgccattaac tttgtcgcac gtctacgatt tgccactgac tttgtcacgt tctacaatat
     1321 gccatcgact ttttcttaac ttctacgatt taccatcgcc gtccggttag ccacttttag
     1381 tactgtacaa atttgttgaa atgaccaaaa tacccctatg acaaaaatat ccaaaatttg
     1441 gataaaatta tcaaaatatt atattataaa cataagattg taaacatcca aaatttgacc
     1501 aaaaacttga aatatgatat ttcataattt ttatccaaat tttgaaacct tttctcccag
     1561 gggtattttg gtcattttga caaatttgta cagtactaac ggagactaac cggacggcga
     1621 tggtaaattg tagaagttaa gcaaaagtcg atggcatatc gtagaacatg ataaagtcag
     1681 tggcaaatca tagacgtgtg acaaagtcag tggtatataa tggtttctct cttttctttt
     1741 tctctcagaa atttgggccc agctaatttt tgccaatgtt gttgcaagaa ctcaacaagt
     1801 gtatcaacct ctcttctgta ttccatctgc cttctatggg tgattattta cacatttata
     1861 tacagggata tacaagagga aaatgcttcg gtgcccctat ccttacgatg tgcctatgtg
     1921 actggccggt gggccagcct tggtcctaat ggtaacttga tggaggacag gagagctgcc
     1981 cccggcatgg gggcatcgct gttcccgagg gaccttgcta ctacccgggt aaaggcatcc
     2041 ccgggtaata gcatgctccg cctagctgcc atcaatagtg tcagtgttag ttggggcgcg
     2101 tagtgcgcat ttattcgcgc ccctagcgcc tggttcttga cttgcttgcg cgagcacatc
     2161 gaggcataat tcactccggt tggcacatcg ggctttaggg gcccctcgga gctccgaagc
     2221 ccttaggtag cctccggagc tcccccttccg gactccgggt acctttcagg cttcggagct
     2281 cggtcactag gtgtgtcacc ggagcccatg gcgctcgacc tcccggagct cttccggagc
     2341 cggtcggccg ccatccgaag cctgcttagg tcttgttcct ctcgccctca atgagatccg
     2401 attctctcgg atctccggag gcctccggag caggggggcc agccgcggac cgacatggcc
     2461 ttgattcacg gtcatctccc ggggaggatg gttctccgga gtgctgcagc ccttcggagg
     2521 ctagcgtccc ttgatccgga gatactcccc taacaccttc ttattcaacc aagctaggac
     2581 caatgaccat gacccttttg gatatcaaga tgaccacagg tttagatatc ctcttaatca
     2641 gccaaccgtt ttccagtcag aaaatcaagt gtgccaacaa gttgcggacc aagaatgttg
     2701 gtggttggtc aggctacatc acttttttctt atatctgtct aagtccatga gctaaaccaa
     2761 aaacatctct cgctcttgct gtcttagctt gcaccgatat tctctgcatc tcggcacgat
     2821 gatatcactc ccattactgc tcttcgtcct cttcttctct gcgctgctgc tcttcccttc
     2881 gagcagtgac gacgacggtg gtggtgatgc tgccggcgac gaactcgcgc tgctctcttt
```

```
2941 caagtcatcc ctgctatacc agggggggca gtcgctggca tcttggaaca cgtccggcca
3001 tggccagcac tgcacatggg tgggtgtcgt gtgcggccgc cggcacccac acagggtggt
3061 gaagctgcgg ctgcgctcct ccaacctggc cgggatcatc tcgccgtcgc tgggcaacct
3121 atccttcctc aggacgctgc aactcagcga caaccacctg tccggcaaga taccccagga
3181 gctcagccgt ctcagcaggc tccagcaact ggtactgaat ttcaacagcc tatcgggtga
3241 gattccagct gctttgggca atctaaccag tctctcggtt cttgagctga ctaacaatac
3301 actgtccgga gcaatccctt catctctggg caaactcaca ggtctcactg atcttgcact
3361 ggctgaaaat acgctgtctg gttccatccc atcatctttc ggccaattgc gcagattatc
3421 tttccttagc ttagccttta acaatttaag tggagcgatc ccagatccta tttggaacat
3481 ctcctctctc accatattcg aagtcatatc caacaagcta agtggtacac tgcctacaaa
3541 tgcattcagt aatcttccta gtctgcagga ggtatacatg tattacaacc agtttcatgg
3601 tcgtatcccg gcatcgatag gtaatgcttc caacatctca atatttacca ttggtttaaa
3661 ctcttttagc ggtgttgttc caccggagat tggaaggatg agaaatcttc agagactaga
3721 gcttccagaa actctttcgg aagctgaaga aacaaatgat tggaaattca tgacggcatt
3781 gacaaattgc tccaatcttc aagaagtgga actgggaggt tgtaaatttg gtggagtcct
3841 ccctgattct gtttccaatc tttcctcttc gcttgtatct ctctccatta gagataacaa
3901 aatttcaggg agcttaccta gagatatcgg taatctcgtt aatttacaat atctttctct
3961 cgctaacaac tccttgacag gatcccttcc ctcttccttc agcaagctta aaaatttacg
4021 tcgtctcact gtagataaca acaagttaat tggttctctc ccattgacca tcggtaatct
4081 tacacaacta actaatatgg aggtccaatt taatgccttc ggtggtacaa taccaagcac
4141 acttggaaac ctgaccaagc tgtttcaaat aaatcttggc cacaataact ttatagggca
4201 aattcccatt gaaatattta gcattcccgc actctctgaa attttggatg tgtcccataa
4261 taacttggag ggatcaatac caaaagaaat agggaaactt aaaaatattg tcgaattcca
4321 tgctgattcg aacaaattat cgggtgagaa ccctagcacc attggtgaat gccaacttct
4381 gcagcatctt ttcctgcaaa acaatttctt aaatggtagc atcccaatag ctctgactca
4441 gttgaaaggt ctggacacac ttgatctctc aggtaacaat ttgtcaggtc agatacctat
4501 gtccttaggg gacatgcctc ttctccactc gctgaacctt tcgttcaaca gcttccacgg
4561 tgaagtgcca accaatggtg tttttgcaaa tgcttctgaa atttacatcc aaggcaatgc
4621 ccatatttgc ggtggcatac ctgaactaca tcttccgacg tgttccttaa aatcaagaaa
4681 gaaaaagaaa catcaaattc tgctgttagt ggttgttatc tgtctcgttt cgacacttgc
4741 cgtcttttcg ttactctaca tgcttctaac ctgtcataag agaagaaaga aagaagtccc
4801 tgcaacgaca tccatgcaag gccacccaat gatcacttac aagcagctgg taaaagcaac
4861 ggatggtttt tcgtccagcc atttgttggg ttctggatct tttggctctg tttacaaagg
4921 agaatttgat agtcaagatg gtgaaatcac aagtcttgtt gccgtgaggg tactaaagct
4981 ggaaactcca aaggcactca agagtttcac ggccgaatgc gaaacactgc gaaatactcg
5041 acaccggaat cttgtcaaga tagttacgat ttgctcgagc atcgataaca gagggaatga
5101 tttcaaagca attgtgtatg acttcatgcc caatggcagt ctggaagatt ggctacaccc
5161 tgaaacaaat gatcaagcag agcaaaggca cttgactctg catcagagag tgtcacgccg
5221 gaatttctat ccaaaattcc aaacgcttac atgtgtgtga accctcgtcc aggaatcagc
5281 cgagacacac aataacaaat tgataataga gtacaattat tactctaatt aataagcgta
5341 taaaatgtca ttacagaggt agatagttcc tctcaatcaa taaagatcta agcagcggaa
5401 aaataagata aacggcgcag acggctccac tccacaggca gcttgaccaa ggctacacct
5461 aatcctccac accatcagct tcactgtaga actcttcctc tgatgaatga ttgcaaggtg
5521 agtatatgac atactcagca agccacgcag caaatatgca agtgcacagg ataacaaagg
5581 atggcatagt agggttttat ttgcaaaagc agcatttagc aaacatttga gaatttaata
5641 aaacagttaa gtaattaaac aatattaatc caacgctata caacataccc tgttgtatag
5701 gcccaaccat tctgaacaac catcccggct gcacagttct atctccaaac caggaatata
5761 ccattccaaa ccaggagcta atcaaattat taccaattaa agcacctttt attatgatga
5821 gaagggtgag actaatcacg aaagatattg ttagacccgc ctataaccgc gggcacggct
5881 attcgaatag ttttactctg atcagaggtg taccactgta cccacaagac acaaccccac
5941 atcatgtcac catgtgcctc aataccacca cggtacctcg gaaaggagtt gtgacaatac
6001 cccttgcata acacaatcca ctgcagtgca ccttcctgga tcataatcac ccccttaaaa
6061 acaaggcatg gactccccag cgaccctcgt gggcttatct ccgccacttc tcagtctggt
6121 gccccgcaat gaaccatgct atataaaaga taaagccgtt gcccatgctg gcttatggtt
6181 ggcacggtta atgtttcaca accgaaactc gtgaaccggt ccttaattgt catgagcacg
6241 accatcaaaa ccatgtgctc acaacccacc attatcaggt tttagttggc aaataattaa
6301 ttaacaaatc acgattgacc atcgtgaact atcattaagc catcattaaa taacagtgag
6361 tcataagtta tcccaatagt aagctaatgt ttctaagcag ggctaagcaa ttatatctaa
```

```
6421 tatctagttg aaccaatata taaagctcac tagtcaaatt ataataaccc aaggtatcaa
6481 ggaataaagt aatcaataac aaaagggcta taacaaacaa taggttaatt ccacccaatg
6541 acattcgaaa ataaatgcaa tatttgaata gaaacaatag ctttaaatag gatcaacatg
6601 ctcaaagggt tgtatgggat ctgtgtgact tgccttgctg gccttggaac tcttcaaact
6661 cttctccggc gaaaacggac tctccggaaa cgacggaatc taaacaaaaa gaagcaaaac
6721 caccaaaaca gcacataaac caactaaatc ggagctaaga tgaattagtt atgaattttt
6781 gaagattaaa tcggattaaa acacttaaat tgattttaat tgaattatga cgcaataatg
6841 aattattttt gaaaaggaaa aggaggatta ttgcgtcagc gggctagggt ttcggtggac
6901 cgggcacaca ggcgacggct cacgcgaacg gacggccgag atcgacttga tccaaaacgg
6961 acggccgaga tcgaacggtc cacgaccggc tcacagcgaa cggccccgat gacgtcggcg
7021 atgacgtcac caccggcggc ggcggctcgg cggctcgggc ttgcacgctc gccggcgaac
7081 gacggcacgg cggcgcgaat ggaaggcacc aacgggtaga gcgcgacgcg gcgaactcac
7141 cggtgaccaa aagagcggcg gaagatcaat ggacggcgac ggcgacgagg aggaagcggc
7201 ggcaaacttc gggtcgacgg tggcgacggt gctccggcgg tcttcggcgg cggcaaagga
7261 gcggacgaga acggcggcga cttggcgatc acgacggtgg ccttcccgag cgatgatgac
7321 gaccgagacg gcggcgacgc acggctggag cgacggctac gacggcggcg ctaggttgca
7381 cggcgctaga gctcttccgg cgacgagagg cgaaggcgaa ggtggcgacg ggtagaggag
7441 acaccgggga acctttaaa ggggctcgca ggcgacggcg aaggcccacg gcggctggcg
7501 acgagaagga aggtttaggg ttcggaggag ggagacgaat ccgattcgaa ctcgattcca
7561 acgatttcca aaacgaatta gccgatgttt ccaaaagaga aaaggtagag gagatcccgg
7621 agattgtttc ccctctatca attcggccgg aaacggaaag gatcgatcga atttggaagg
7681 gaacggcggc ggcgcgaaac tagggtttcg ggcggcggcg gccggaggtt gacgacgacc
7741 ctgacaggtt ggccccacct gtcagcgggc ggacgcgcgc gcgcggcggc ggactgggcc
7801 ggactgggcc gaggagagag agagcggttt tgggccgact ttcggcccaa agccaaaaga
7861 gacttttaa aacctttttc aatttaaatt attcatgaaa tgtaattcca tttattaaaa
7921 atacttcctt agctcaaata aatcccagaa aaatctagga attatagaat taagcaaagt
7981 atttaatgaa attttatctg gccccatttt atattgtaat ttattaattt aaaattagat
8041 cttctcttct aggcttttaa aataaattct aaaaattcca tttaaacaac aatttatata
8101 ttttgaattt tcagggtgtg acagagagtg cccatactac gtgatgttgc atgtgcattg
8161 gaccatcttc acttccatgg ccctgaccct attgtacact gtgatattaa atcaagcaat
8221 gtgttgttag atgctgatat ggtagcccat gttggagact ttggacttgc aagaatactt
8281 attgagggaa gctcattgat gcaacagtca acaagttcga tgggaatcag ggggacaatt
8341 ggttacgcag caccaggtta atcctaaact gtttatgtct acctcctttc attgtttttt
8401 ttagatttgc tctggtccaa caaaaaatac ctaaagatac agatacttgt acctcacagt
8461 actaaatagt tttcgatcat tgcattgtta gatccaacga tcaggaaacg atttggtacc
8521 gtgaccgtga ggtatcggaa tctcgagata ttttttgttc gaccgtagca aatctatttt
8581 tttgtttgtt ttcttctctt taatgtttta tgactatgaa ataattttta tttctggaaa
8641 acagagtatg gtgtcgggaa cactgcctcg acacatggag atatttacag ttatggaatt
8701 ctagtgttgg aaacagtaac cgggctgcgg ccggcagata gtacattcag acctggcttg
8761 agcctccgtc agtacgttga accgggtcta catggtagac tgatggatgt tgttgacagg
8821 cagcttggtt tggcttccga gacatggctt caggctcgag atgtttcgcc atgcagcagt
8881 attactgact gccttgtttc actgcttaga cttgggctgt cttgctctca ggaattgcca
8941 tcgagtagaa cgcaagccgg agatgtcatc aatgaactgc gtgccatcac agcgtctctc
9001 tcgatgtcat ccgacatgtg aagatgtgag acatgctgat gttatgtccg agtatttcgt
9061 tgtaatgtaa tgtgaagggt gagtgtgtga ctgcttggtt gtaagctatt tcctgatctg
9121 cccatcagat catgtatctg ttctattgtt gtatttctca gaacaactac acaccctaag
9181 taggagtaca caatagtgta tttgtgtgat ttcaatattg atgcataccc atgctatgtg
9241 ctaaaattat atactgaaat tttgagatgt ctgaagttaa cagtcaatcg gggagcgatt
9301 cacaccatac cgcgaaatcg acctaatcag ctaatctaat tctacaggct gcctttgcat
9361 gacagtgtga tattaaatta gcccagccct ttttagcaaa cgatgggagg gtcaatgctc
9421 taga
```

SEQ ID NO:9

/translation="MISLPLLLFVLFFSALLLFPSSSDDDGGGDAAGDELALLSFKSS

LLYQGGQSLASWNTSGHGQHCTWVGVVCGRRHPHRVVKLRLRSSNLAGIISPSLGNLS

FLRTLQLSDNHLSGKIPQELSRLSRLQQLVLNFNSLSGEIPAALGNLTSLSVLELTNN

TLSGAIPSSLGKLTGLTDLALAENTLSGSIPSSFGQLRRLSFLSLAFNNLSGAIPDPI

WNISSLTIFEVISNKLSGTLPTNAFSNLPSLQEVYMYYNQFHGRIPASIGNASNISIF

TIGLNSFSGVVPPEIGRMRNLQRLELPETLSEAEETNDWKFMTALTNCSNLQEVELGG

CKFGGVLPDSVSNLSSSLVSLSIRDNKISGSLPRDIGNLVNLQYLSLANNSLTGSLPS

SFSKLKNLRRLTVDNNKLIGSLPLTIGNLTQLTNMEVQFNAFGGTIPSTLGNLTKLFQ

INLGHNNFIGQIPIEIFSIPALSEILDVSHNNLEGSIPKEIGKLKNIVEFHADSNKLS

GENPSTIGECQLLQHLFLQNNFLNGSIPIALTQLKGLDTLDLSGNNLSGQIPMSLGDM

PLLHSLNLSFNSFHGEVPTNGVFANASEIYIQGNAHICGGIPELHLPTCSLKSRKKKK

HQILLLVVVICLVSTLAVFSLLYMLLTCHKRRKKEVPATTSMQGHPMITYKQLVKATD

GFSSSHLLGSGSFGSVYKGEFDSQDGEITSLVAVRVLKLETPKALKSFTAECETLRNT

RHRNLVKIVTICSSIDNRGNDFKAIVYDFMPNGSLEDWLHPETNDQAEQRHLTLHQRV

SRRNFYPKFQTLTCV"

SEQ ID NO:10

DEFINITION Oryza longistaminata receptor kinase-like protein, family member A2, pseudogene sequence.
ACCESSION U72727
SOURCE long-staminate rice.
ORGANISM Oryza longistaminata
Eukaryotae; mitochondrial eukaryotes; Viridiplantae; Charophyta/Embryophyta group; Embryophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.

FEATURES	Location/Qualifiers
source	1..5940
	/organism="Oryza longistaminata"
	/strain="IRBB21"
	/chromosome="11"
	/map="11q, RG103"
CDS	2151..5855
	/note="family member A2; receptor kinase-like protein truncated by two mutations"
	/codon_start=1
	/pseudo
mutation	2501..2503
	/note="mutation found when compared to family member A1,GenBank Accession Number U72725"
	/citation=[1]
	/replace="acc"
mutation	4355..4356
	/note="mutation found when compared to family member A1,GenBank Accession Number U72725"
	/citation=[1]

```
                     /replace="ac"
     3' flanking     5249..5940
     misc_feature    5453..5697
                     /note="Tourist-Ol2, transposon-like element"
BASE COUNT     1570 a   1200 c   1188 g   1982 t
ORIGIN
        1 tctagaaatt gaattttaaa attgaatata atagagaaat acaactgaac ataaatattt
       61 tgtcattttc attccattta catgttttta aatctaaatt tacagtgcat tttgtatgta
      121 attctaatac aaatttactg taatttgtat ttttagtcat tttttcagtt tctaaaagta
      181 cagtgcactt cagtgacctc cttagtaagc tatgggggga aaatggatat gcacatctga
      241 tctatctatg aaacatagaa gaaatattca tcaccatgta ctacttttta tgaatgtttg
      301 ttggccagat aggatctgaa agatttctag tacagggttg tttcatgtat tgtttcagat
      361 aggatctgaa ataaccactg tacagggttg tttcttgtat ggtttccgca tatttaaact
      421 tttggttgca gatgtatggc taatcaatta tgcagtttta tggcaaacaa aaagggtggc
      481 aatagagtat atttcatcat gtggtatatg attgtttttt gttgattaga ccagaatagt
      541 tgtgtgtatt tgtttttttg ttctgagaac tattgtctgt tctttatttt cattcagaca
      601 atatgaatgc tgttttgtac agattctttt tgttgtgatg ttcagacaat atgaatcagt
      661 ggctgtttgt ctgtttttca ttcaaaattt tcagagtgca caagtgtgtg ttgttaatcc
      721 agaaattttc agaggattgc ttgttctgtt ctgacctgaa ctagtaaaaa aacaaattca
      781 gatttggttc aatattcaga ttcttgtttt tgtaatcatt gctgtgcaga atctgtttgt
      841 actcctggtc atttagattt ttagtctgtg tgtttgttgt tcactgagta tttccattgg
      901 gggcgagagt ctcttttctt ctatgctttt atacagaggt tccaaatcaa ttgcagtgtt
      961 gtgcaatatt cagtgttgtt cagatctatg tatatagctt ctgttctatt ataatgcctg
     1021 cagtaatctg ctcatgtaaa taacagaaac cctgtgggta ttttgcatat ttaggtacaa
     1081 aattcaaagt ttgcagaaac aaatctgttg agcattatgt cttctaaacc actaaacagt
     1141 tttttttctg tgctgatgtt ttttttctgc cctttatgtg aatgtttact tctctctaat
     1201 ctctgtttgg ttttaattca tcagaggagc tattttgcaa gaataaatgt caaagaacaa
     1261 catgaagaaa aggaagactc aaaaatatgc tgcgggtcac aggtaccttt tgatttgtga
     1321 ttttcttgaa ttttgatccc caatatttct tgttttgagt ctgatctgga agaatcttgc
     1381 acttttgatc cttctccaat tctgagtttc ttttccttca tctccatctt tttccccagt
     1441 tttctttctt ctcttcgatc gatctgagtt tacctcttct tctctgtgtt tgtctggaag
     1501 aattattttg cctgttttga ttactcgtca cgtgaatatg tcagcgatct tcttgcattg
     1561 caagttttgt tgtcacgtgg gtttgctccc ttctttttgg gcttcatttg ggctttttct
     1621 tgagctcata atctgatttt ttttttcttt tcttgtttac taattgagcc taagttgggc
     1681 cctgaaagat tcgtttttta gttctgttgg tgtcaaaaat ctgggctgtt ggatcaaaat
     1741 ctgttggctg aaaaattcga cagatttgga actctatttt ctgttgactg aaagattagt
     1801 tttttagttc tgctggagtc aaagtctggg ctgttggatc aaaatctgtt ggctgaaaaa
     1861 ttcgacagat ttggaactct attttctgtt gactgaaaga ttagtttttt agttctgctg
     1921 gagtcaaagt ctgggctgtt ggatcaaaat ctgttggctg aaaaattcaa cagattagga
     1981 actctatttt ctgtcgactg aacagtagac agagttaagc agaaacgaat atcacaattg
     2041 ctatgttcat tgtcttgcgt gagcgctttt tcttctatct gtctgtctag tgcatgagct
     2101 aaaccaaaca tctctcgctc ttgcaccaat attctctgca tctctgcaca atgatatcac
     2161 tcccgttatt gctcttcgtc ctgttgttct ctgcgctgct gctctgccct tcgagcagcg
     2221 atgatggtga tgctgccggc gacgaacttg cgctgctctc tttcaagtca tccctgcgat
     2281 accaggggggg cttgtcgctg gcatcttgga acacgtccgg ccacggccag cagcactgca
     2341 catgggtggg tgttgtgtgc ggccgccggc acccacacag ggtggtcgag ctgcggctga
     2401 actcgtccga cctgtccggg atcatctcgc cgtcgctggg caacctgtcc ttcctcagga
     2461 cgctggacct cagcgacaac cacctgtccg gcaagatacc ctaggaactc agcagtctca
     2521 gcaggctcca acaactggta ctgaatttca acagcctatc gggtgagatt ccagctgctt
     2581 tgggcaatct aaccagtctc tcggttcttg agctgactaa caatacactg tctggagcaa
     2641 tcccttcatc tctgggcaaa ctcaccagcc tcactgatct tgcactggct gaaaatatgc
     2701 tgtctagttc catcccttca tctttcggcc aattgcgcag attatctttc cttagcttag
     2761 cctttaacaa tttaagtgga gcgatcccag atcctatttg gaacatctcc tctctcacca
     2821 tattcgaagt catatccaac aagctaagtg gtacactgcc tacaaatgca ttcagtaatc
     2881 ttcctagtct gcaggaggta tacatgtatt acaaccagtt tcatggtcgt atcccggcat
     2941 cgataggtaa tgcttccaac atctcaatat ttaccattgg ttttaactct tttagcggtg
     3001 ttgttccacc ggagattgga agcatgagaa atcttcagag actagagctt ccagaaactc
     3061 ttttggaagc taaagaaaca aatgattgga aattcatgac ggcattgaca aattgctcca
```

```
3121 atctacaaga agtggaactg ggaggttgta aatttggtgg agtcctccct gattctgttt
3181 ccaatctttc ctcttcgctt gtatctctct ccattagaga taacaaaatt tcagggagct
3241 tacctagaga tatcggtaat ctcgttaatt tacaatatct ttctctcgct aataactcct
3301 tgacaggatc ccttccctct tccttcagca agcttaaaaa tttacgtcgt ctcactgtag
3361 ataacaacaa gttaattggt tctctcccat tgactatcgg taatcttaca caactaacta
3421 atatggaggt ccaatttaat gccttcggtg gtacaatacc aagcacactt ggaaacctga
3481 ccaagctgtt tcaaataaat cttggccaca ataactttat agggcaaatt cccattgaaa
3541 tatttagcat tcccgcactc tctgaaattt tggatgtgtc ccataataac ttggagggat
3601 caataccaaa agaaataggg aaacttaaaa atattgtcga attccatgct gattcgaaaa
3661 aattatcggg tgagatccct agcaccattg gtgaatgcca acttctgcag catcttttcc
3721 tgcaaaacaa tttcttaaat ggtagcatcc caatagctct gactcagttg aaaggtctgg
3781 acacacttga tctctcaggt aacaatttgt caggtcagat acctatgtcc ttaggggaca
3841 tgcctctgct ccactcgctg aacctttcgt tcaacagctt ccacggtgaa gtgccaacca
3901 atggtgtttt tgcaaatgct tctgaaattt acatccaagg caatgccctt atttgcggtg
3961 gcatacctga actacatctt ccgacgtgtt ccttaaaatc aagaaagaaa aagaaacatc
4021 aaattctgct gttagtggtt gttatctgtc tcgtttcgac acttgccgta ttttcgctac
4081 tctacatgct tctaacctgt cataagagaa taaagaaaga agtccctaca acgacatcca
4141 tgcaaggcca cccaatgatc acttataagc agctggtaaa agcaacagat ggtttttcgt
4201 caaccaattt ggtgggctct ggatcgtttg gctctgttta cagaggagaa tttgatagcc
4261 aagatggtga aagcccaaga cttgtcgccg tgaaggtact aaagctggaa actccaaagg
4321 cactcaagag tttcacggcc gaatgcgaaa cactgtgaaa cactcgacac cgcaatcttg
4381 tcaagatagt tacaatttgc tcgagcatcg ataacagagg gaatgatttc aaagcaattg
4441 tgtatgactt catgcccaat ggcaatctgg aagattggct acaccctgaa acaaatgatc
4501 aagcagagca aaggcacttg actctgcatc agagagtgac catactactt gatgttgcct
4561 gtgcattgca ctatcttcac cgccatggcc ctgaacctgt tgtacactgc gatattaaat
4621 caagcaatgt gctgttagat gctgatatgg tagcccatgt tggagacttt ggacttgcaa
4681 gagtacttat tgagggaagc tcattgatgc aacagtcaac aagttcgatg gggataaggg
4741 gaacaattgg ttacgcagca ccaggttaag tctaaactgt ttatgtctac ttcctataat
4801 cttctctttt ttgaggtttc ttctctctag tgttttatga ctatgaaata ttttttgcta
4861 ctggaaaaca aagtatggtg tcgggaacac tgcctcgaca cctggagata tttacagtta
4921 tggaattcta gtgttgaaaa cagtaaccgg gaagcggccg acagatagta cattcagaac
4981 tggattgagc ctccgtcagt acgttgaacc gggtctacat ggtagactaa tggatgttgt
5041 tgacaggaag cttggtttgg attccgagaa atggcttcag gctcgagatg tttcgccatg
5101 cagcagtatt agtgaatgcc ttgtttcact gcttagactt gggttgtctt gctctcagga
5161 attgccatcg agtagaatgc aagccggaga tgtcatcaat gaactgcgtg ccatcaaaga
5221 gtccctctcg atgtcatccg gcatgtgaag atgttggagt atttcattgt aatgtgatgt
5281 gtctatcagt acccttcaca actgatttca ttctgccgtg gtatttagtt atttacaaga
5341 gagtcactga agggtgagtg tgtgactgct tggttgtagc tatttcctga tctgcccatc
5401 agatcatgta tctgttctat tgttgtattt ctcataataa ccacacacct aagggagggt
5461 tcggcagagg agattgtgag ttagtttgtt ttgttttcca cgcgcacgct tcccgaacta
5521 ctaaacggtg tgtttttttgc aaaaaaattt ctatatgaaa gttgctttaa aaaatcatat
5581 taatccattt ttgaagttta aaatagttta tactcaatta atcatgtact aatggctcac
5641 ctcgttttgt gtatcttccc aatcttctct tttcccctcc tctcaaactc accctaagta
5701 cacaacactg tatttgtgtg atttcaatat tgatgcatat atacccatgc tatatgctag
5761 aattatatac aaaaattttg agatgtctga agttaacaat caatcaggga gcgattcaca
5821 ccaaaccgcg aaatcgacct aatgagctaa tctaattgta caggctgcct ttgcatgaca
5881 gtgcgatatt aaataagccc agccctttt agcaaaggat gggagggtca atgttctaga
//
```

SEQ ID NO:11

ORGANISM Oryza longistaminata
Eukaryotae; mitochondrial eukaryotes; Viridiplantae;
Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
Liliopsida; Poales; Poaceae; Oryza.

FEATURES	Location/Qualifiers
source	1..7204
	/organism="Oryza longistaminata"
	/strain="IRBB21"

```
                     /chromosome="11"
                     /map="11q, RG103"
     CDS             join(1683..1758,1771..2186,2347..4352,5147..5547)
                     /gene="Xa21"
                     /note="disease resistance gene Xa21 family member F"
                     /codon_start=1
                     /exception="Author-given protein sequence differs from
the
                     conceptual translation for reasons explained in
citation."
                     /pseudo
                     /product="receptor-like kinase protein"
     flanking_region 1..1682, 5548..7204
     misc_feature    233..384
                     /note="Gaigin-011, transposon-like element"
                     404..532
                     /note="Gaigin-012, transposon-like element"
                     1080..1239
                     /note="Tourist-011, transposon-like element"
                     6201..6583
                     /note="Crackle, transposon-like element"
                     6750..6956
                     /note="Ds-rice3, transposon-like element"
     repeat_region   7053..>7204
                     /note="number of AT repeats has not been determined; 3'
                     flanking sequence found in GenBank Accession Number
                     U72729"
                     /rpt_family="microsatellite"
                     /rpt_type=tandem
                     /rpt_unit=7053..7054
BASE COUNT     2124 a    1455 c    1309 g    2316 t
ORIGIN
        1 aagctttcta aattatttaa ctctaagtct gttattatcc ccaagtacat catcatcata
       61 cataatattt catattcacg acatcccttaa gctagatgct tttggccatt ctcttatctt
      121 tttaaagaaa ttctctcccca attaagatga gagtgtcttc tagcaatttg ccagttttta
      181 caatgtcttt gagtcctcac acatttttcat gatgttacca ataaattacg aacgccgtgt
      241 ttagttctaa agtttttctt caaacttaca acttttcaat cgcatcaaaa ctttctccta
      301 cacacacaaa ctttcaactt ttccatcaca tcgttccaat ttcaaccaaa cttccaattt
      361 tggtatgaac taaacacagc cgaaaacaaa atttgtgtgt tatggccctg tttagattct
      421 aacttttcca ttacatcaaa ctttcctaca tacacgaact ttcaactttt ccgtcacatc
      481 gtttcaattt tttaaaactt ccaattttaa cgtggaacta aacacaacct atataacgaa
      541 atttgtcaaa aacttaatgg tgaaagtcac acctcaaagg aagggcgcgc ctctagtcaa
      601 gaacatcaat taaaaaggta cacaggttgt actagcttgt tcatgtttaa tcttgcgtct
      661 gcgagacgct aaatccatgc caaacaaaag tgcttctata gagataatca taagaatatg
      721 gtttgggacc atatccaact gctcagaaga atctcgttcg gaggtgaagg ttaagatgtt
      781 cacctctcca cacataaaac aaagcgatct ttttcgcata attaattaag cattagataa
      841 aataaactta aaaaataaat caatatgatt tttttagaaa aaaaatatat acactaagta
      901 taagcattgt caaggaggaa gaaacacaca cttccatata gagagataga aacatagcta
      961 taggtagtgt cactgagtat ttttcatcac gcatatgcat ataaaattag ggggtgttta
     1021 catccatagg tgtaaagttt tggcatgtta tatcgagtat tacgtagaat gtcgtattag
     1081 gtgttcgggc actaataaaa aaataattac agaatccgtt agtaaaccgc gagataaatt
     1141 tattaagcct aattaatcca tcattaacaa atgtttaccg tagcaccaca ttgtcaaatc
     1201 atggagcaat taggtttaaa agattcgtct cgcaaattag tcataatctg tgcaattagt
     1261 tatttttaga ctatatttaa gacttcgtac aggtgttcaa acgttcgatg tgacatggtg
     1321 caaaatttta gggtgtcatc tagacactcc cttaattaga aagttaggaa gaggcggtaa
     1381 agaacgcagc atgactgaaa ctttgaaaat ttgataaggt acaccaactg gagtatcttt
     1441 tattttcatt gaagactttg accagaagag cttgacccgt tttcttgga gtagccagta
     1501 atgtttcatt ctttttcttt tgctgggact tctttttatt ttttttgaca ggagccattt
```

```
1561 gttgggactt gggatccctt tactgttata ggaccagtgc ttgaatccaa acactgcatt
1621 gatcagctca gctcattgta gcgcactcct ccgcatgcat ggcgagatca ccaacgtcgg
1681 tcatgatctc ttctttgctg ctgctgctgt tgatcggccc agcgagcagt gacgatgatg
1741 ctgctgctgc tgctgctcgt accagtacag gcggcgtcgc cggcgacgaa ctcgcgctgc
1801 tctctttcaa gtcatccctg ctacaccagg ggggcttgtc gctggcatct tggaacacgt
1861 ccggccacgg ccagcactgc acatgggtgg gtgttgtgtg cggccgccgc cgccgccggc
1921 acccacacag ggtggtgaag ctgctgctgc gctcgtccaa cctgtccggg atcatctcgc
1981 cgtcgctggg caacctgtcc ttcctcaggg agctggacct cagcgacaac tacctctccg
2041 gcgagatacc accggagctc agccgtctca gcaggcttca gctgctggag ctgagcggta
2101 actccatcca agggagcatc cccgcggcca ttggagcatg caccaagttg acatcgctag
2161 acctcagcca caaccaactg agattggtgc cagcttgaaa catctctcga atttgtacct
2221 tcacaccaat ggtttgtcag gagagattcc atctgctttg ggcaatctca ctagccttca
2281 gtattttgat ttgagctgca acagattatc aggagctata ccttcatcgc tagggcagct
2341 cagcagcagt ctattgacta tgaatttgcg acagaacaat ctaagtggga tgatccccaa
2401 ttctatctgg aacctttcgt ctctaagagc gtttagtgtc agcgaaaaca agctaggtgg
2461 tatgatccct acaaatgcat tcaaaaccct tcacctcctc gaggtgatag atatggacac
2521 taaccgtttc catggcaaaa tccctgcctc agttgctaat gcttctcatc tgacacggct
2581 tcagattgat ggcaacttgt tcagtggaat tatcacctcg gggtttggaa ggttaagaaa
2641 tctcacaaca ctgtatctct ggagaaattt gtttcaaact agagaacaag aagattgggg
2701 gttcatttct gacctaacaa attgctccaa attacaaaca ttggacttgg gagaaaataa
2761 cctgggggga gttcttccta attcgttttc caatctttcc acttcgctta gttttcttgc
2821 acttgatttg aataagatca caggaagcat tccaaaggat attggcaatc ttattggctt
2881 acaacatctc tatctctgca acaacaattt cagagggtca cttccatcat cgttgggcag
2941 gcttagaaac ttaggcattc tagtcgccta cgaaaacaac ttgagcggtt cgatcccatt
3001 ggccatagga aatcttactg aacttaatat cttactgctc ggcaccaaca aattcagtgg
3061 ttggatacca tacacactct caaacctcac aaacttgttg tcattaggcc tttcaactaa
3121 taaccttagt ggtccaatac ccagtgaatt attcaatatt caaacactat caataatgat
3181 caatgtatca aaaaataact tggagggatc aataccacaa gaaatagggc atctcaaaaa
3241 tctagtagaa tttcatgcag aatcgaatag attatcaggt aaaatcccta acacgcttgg
3301 tgattgccag ctcttacggt atctttatct gcaaaataat ttgttatctg gtagcatccc
3361 atcagccttg ggtcagctga aaggtctcga aactcttgat ctctcaagca acaatttgtc
3421 aggccagata cccacatcct tagcagatat tacttatgctt cattccttga acctttcttt
3481 caacagcttt gtggggaag tgccaaccat tggtgctttc gcagatgcat ccgggatctc
3541 aatccaaggc aatgccaaac tctgtggtgg aatacctgat ctacatctgc ctcgatgttg
3601 tccattacta gagaacagaa agcattttcc agctctacct atttctgttt ctctggtcgc
3661 agcactggcc atcctctcat cactctactt gcttataacc tggaacaaga gaactaaaaa
3721 gggagcccct tcaagaactt ccatgaaagg ccacccattg gtctcttatt cgcagttggt
3781 aaaagcaaca gatggtttcg cgccgaccaa tttgttgggt tctggatcat ttggctcagt
3841 atacaaagga aagcttaata tccaagatca tgttgcagtg aaggtactaa agcttgaaaa
3901 tcctaaggca ctcaagagtt tcactgccga atgtgaagca ctacgaaata tgcgacatcg
3961 aaatcttgtc aagatagtta caatttgctc gagcattgat aacagaggga acgatttcaa
4021 agcaattgtg tatgacttca tgcccaacgg cagtctggaa gattggatac accctgaaac
4081 aaatgatcaa gcagaccaga ggcacttgaa tctgcatcga agagtgacca tactacttga
4141 tgttgcctgt gcattggact atcttcaccg ccatggccct gaacctgttg tacactgtga
4201 tgttaaatca agcaatgtgc tgttagattc tgatatggta gcgcatgttg gagattctgg
4261 gcttgcaaga atacttgttg atgggacctc attgatacaa cagtcaacaa gctcgatggg
4321 atttagaggg acaattggct atgcagcacc aggtcagcaa gtccttccag tattttgcat
4381 tttctgatct ctagtgctat atgaaatagt ttttacctct agtgaaactg atggagaata
4441 taagtaatta attgaactaa ttaaattgca caaaaataag attatttgcc atatctattc
4501 agatgctaaa tatagctagt tcatagaggt acatattttt tttatatagg aatctagagc
4561 tactacacac tcaaatcaaa ttatgggtgt tttctgctct acactgcaat atgaaatgat
4621 tatcagaagg atcaaatttg agtaaatttg tcaattctac atttaagaaa cacttttttt
4681 tgtatgtact agttattaca atttttttatt tcaagaactt gcattgacca tgaaaagtac
4741 ttggtactac ttctaattcc cacatggagg tggtgaaaat aatatagata caaaaacgaa
4801 gtatcatatg ttgtgtgata tactataatc acaatgaaca caaacaggat tcgtacaaaa
4861 gtaattggcc atcatagcaa ctgattgctt ggggtaactg tatagcacaa tcataccaaa
4921 tttctttaga tatgtatttg taaattagat tcttaaagtt aaatatgaaa tttcattggt
4981 atttatgttt ctttatataa taaaaattaa tccaaccttt acatctacca tttgtccagc
```

```
        5041 catccttgtt atttgtgata tttaacacgt aattttacat aattatacat ccaagttctt
        5101 tttatttaac actggaaatt tgaaatcgta ttcctactc aaacagagta tggcgtcggg
        5161 cacattgcat caacacatgg agatatttac agctatggaa ttctagtgct ggaaatagta
        5221 accgggaagc ggccaactga cagtacattc agacccgatt tgggcctccg tcagtacgtt
        5281 gaactgggcc tacatggcag agtgacggat gttgttgaca cgaagctcat tttggattct
        5341 gagaactggc tgaacagtac aaataattct ccatgtagaa gaatcactga atgcattgtt
        5401 tcgctgctta gacttgggtt gtcttgctct caggatttgc cattgagtag aacgccaacc
        5461 ggagatatca tcgacgaact gaatgccatc aaacagaatc tctccggatt gtttccagtg
        5521 tgtgaaggtg cgagcctcga attctgatgt tatgtcttgt aatgttttat tgccactagt
        5581 cttcagattg gaatgctctt ccgatcagac ttcttcagtg gtatctacca cacgatcact
        5641 aaagtcatcg tggctatttc ctgatccagc atatctgatc atgcatgttc tgtgttttat
        5701 acctgtattt tactctgaat tgccacacct caaccctgcc tctgtttgtt tggcatacaa
        5761 aagatagtga tgagtatatt gtttcagggg cttcctagtt ggcgtgtgtg cttaccggca
        5821 cgcacgcagc ccgagggtgg gtttcttttt ttttccattg ttattccgtt gcttttttcc
        5881 accacggtag attttttttt tctggatttc catttttcc gttgtttttc tctatcgctt
        5941 atgctggcgg attttttttcc gtggtttttt tttcaagacg agtatatcta atgtaactaa
        6001 catgttactt ttagataacg atggttatta agataagatt tttttctgga agatttttgt
        6061 aagtaaatgg taaaaaatat ggaaatggaa acggaaatag ttttgctgtt ataccgatcg
        6121 tttccatatt taccgtattc ttatagaaat taccgtttct tataatatgg taattaccgt
        6181 atttctaaat atgttgatat cgatttttgct atatattgcg acaaattttc tcccaaaaat
        6241 ttgatagatg taattatagt acaatcgtag tgtaattaca ctgtaactat agtgtaactt
        6301 gtatgtaact ttcaaaaatc tctctgtaat atgttatttt ggtaaaatag aggttgtggg
        6361 aacaaatcct tacacatatg tgtgcgctgt gatttttctt cttcctcacc aaaacaaaac
        6421 ttataataga tttaacaatt caaaattacg tgaaacttat acaagttaca ccgtagttac
        6481 atgcaagtta cagtgtaatt acactacgat tgtactataa ttacatctgt caaatttttа
        6541 ggagaaaatt tgtcgacaaa tatataggtg atcccgttga tatttatagg gtatgtctct
        6601 acttgactca cagtttagag attgattgac tatttaatca aatccctaac ttgattgcac
        6661 ggctaaaatg gagttgattt ctaatttata tagtatagat tgaatttatt cgtacatata
        6721 acatacttat gtaaagttaa atatatgttt tctatagttt aatgtttctg tatttgttac
        6781 cggttttcga tctataccga ccatgtttcc ttcagtatta ttccgtttcc ggttttctga
        6841 tatttctgat atcgttttcg tttccgagtt taccgttttc gatttcattt ccgagaaaaa
        6901 tatgattatg gaaatggttg aggctgtttt ccgatcgttt ccgaccgttt tcttccctac
        6961 ccgtagcaat aatatataat attttatctc taatctttct ctctctcata tcaacgaata
        7021 ttcgctaaga gactgctatt aacaaggctt ttatatatat atatgtacat atatatatat
        7081 atatatatat atatatagac acacacacat acatacatac atacatacat atatatacat
        7141 acatatacat atatatatat gtatacatac atatacatat atatatatgt atacatacat
        7201 atac
//
SEQ ID NO:12
DEFINITION  Oryza longistaminata receptor kinase-like protein, 3' flanking
sequence of
            family member F.
ACCESSION   U72729
SOURCE      long-staminate rice.
  ORGANISM  Oryza longistaminata
            Eukaryotae; mitochondrial eukaryotes; Viridiplantae;
            Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
            Liliopsida; Poales; Poaceae; Oryza.
FEATURES             Location/Qualifiers
     source          1..1332
                     /organism="Oryza longistaminata"
                     /strain="IRBB21"
                     /chromosome="11"
                     /map="11q, RG103"
     gene            join(U72728:1683..7204,1..1332)
                     /gene="Xa21"
     misc_feature    1..1332
                     /gene="Xa21"
```

/note="sequence 3' of the microsatellite sequence present in GenBank Accession Number U72728"
BASE COUNT 233 a 495 c 380 g 224 t
ORIGIN

```
        1 gccgtcgatc aatcatttgg aaacggccca catcttttcc atctatatgc attcatgaaa
       61 tacatggtat atcccatcga tcggacatca cctgttagcg cgtacgccat cgtcgtgatc
      121 aacctagcta gggcaaacgt cctccgatcg ccaccatcac caatgaacaa gctgctgcgg
      181 cctctcggtg gcctgaggtt gctcaaccga gaagaacatc cgttccgatg cttctcctcc
      241 tccatcgatc tcgtcttccc aggtcgccgc cgccgccaca tggcaaccac cgtgacccac
      301 ccgccgccga cggaatcccg ctggttcgac ggcggcggcc gcgactgctg acccggcctc
      361 ggtgatgctg gaacgttggg gctgcctcag gggctccacg ccggcgaacg tcgccgccga
      421 cgacaacacc gccgcggagt cccgcacctc ccgcggccaa ccctccgccg tcgccctcgc
      481 ccgcgcgtcg ccgccggcga tctccttcat ctgcttcgat cgcggggatg atggctacgt
      541 catcgcggct cacggcgact ctgtcctctt ccggatgagt tggaacgact acttcgtcta
      601 catggccgcc gccggcaagc cgccgtcgct gacgctgctc cccgtctgcg acatccccat
      661 gaacgagcgc tgctgggtca gcaaggaccg tttcaacgac agcttccgca ccacgggccg
      721 ggtgttcgac cagcaggaca ccggcatcct gcgcctccgc ggcggcgagg aggcgccgcc
      781 tctagtggcg cagctccaga tcgcgcacga ggcgccgttc gacacggccg agctctgcgt
      841 gctccgcccc ggccacggcc acggcgagtg ggagctcaag acggcggtgc ccatcgtcca
      901 ccacgacggc ggcggcgaac gccgccatgg cctggagatg tggcaggaga caacgtggcc
      961 gtccccgtcg gcgaccgctt catgtgctgg gccaactacg acctcgccac cttcctcatc
     1021 tgcgacatgg cggcggcgga tctcgacaac cccaagctcc tgtacgttcc gctgccggtg
     1081 aaccagtgcc acccaaggga gagcgacttc gacgacgacc accaccacga cgagctgatt
     1141 ccatggggag tacttccgca acatcgtcgc caccggcgcc gacggcggcg acgacattgt
     1201 gcgattcgtc agcatccaca accgctgctg ctgcggcgcg cccgtgatac acaacctgtg
     1261 cgaacgctcc agctcggcgt tcatggtgaa catctggaag cttgcctgcg gaaccccccgc
     1321 gcccgcgaca gc
```

SEQ ID NO: 13

```
TCAAGCAACAATTTGTCAGGTCAAATCCCTGAATTTTTGGCAGGGTTTAGTTTTATATATCTTAACTTATCTTTCA
ATAATTTTGAAGGCAGAGTGCCAACAGATGGGATATTCAAGAATGCAAGCATTGTTTCAGTCACAGGAAACTCTAA
GCTTTGTGGAGGCATACCTGAGTTTCAACAGCCTGCATGCAACTTCAAAAGGTCTGAAAAAAGGAGAGTTAAGGTA
ATTGTTGGTATTATTGCAGGAGGTTTAGGAGCAATTTTGGTGGTGTTGTCCTTTATATTTCTTTTGAGATTAAGAA
AGAAAAGTCACAAACCCAGTTCATCCTATTCAGAAAATTCACTTTTGGAACTTCCAAAAGTGTCATATAGAGATCT
CTATAAGGCCACTGATGGGTTCTCCTCAGAAAATTTAATTGGTACTGGTAGTTTTGGGTCCGTATATAAAGGAATT
CTTGATGAAGGTGGACCAGTTGTTGCTGTTAAAGTGCTTAACCTCCAGCATCATGGAGCAGCTAAGTCTTTCATGG
CTGAATGTGAAGCCTTGAGAAATATCAGACACCGGAATCTTGTAAAGATACTAACTGCTTGTTCAGGTGTTGATTA
TCAAGGCAATGATTTCAAGGCACTGGTTTATGAGTACATGGATAATGGAAACCTTGAGGAGTGGTTGCATCTACCA
GTTTCAGCAGATAGAAATCATGGGGAGCCTAAGAATCTAAATCTTCTTCAGAGAGTAAATATTGCAATTGATGTTG
CTTCTGCAATTGAATATCTCCATCATCATTGCGGAAATCCAATAGTTCATTGTGACCTTAAATCAAGCAATGTGCT
GTTA
```

SEQ ID NO: 14

translation (1st frame):

```
SSNNLSGQIPEFLAGFSFIYLNLSFNNFEGRVPTDGIFKNASIVSVTGNSKLCGGIPEFQQPACNFKRSEKRRVKV
IVGIIAGGLGAILVVLSFIFLLRLRKKSHKPSSSYSENSLLELPKVSYRDLYKATDGFSSENLIGTGSFGSVYKGI
LDEGGPVVAVKVLNLQHHGAAKSFMAECEALRNIRHRNLVKILTACSGVDYQGNDFKALVYEYMDNGNLEEWLHLP
VSADRNHGEPKNLNLLQRVNIAIDVASAIEYLHHHCGNPIVHCDLKSSNVLL
```

SEQ ID NO: 15

DT4

```
GGATCCCCGCGCAGCTTGGCCGGTCGTCTTCGCTCCAGCGCGTGCGCCTGGGAAGCAACATGCTCTCCGGGCCGAT
CCCGCCGTCGCTTGGCGGAATCGCCACGCTGACCCTGCTGGATGTGTCAAGCAACGAGCTCACGGTGGCATCCCAG
CGGCGCTCGCTCAATGCAGGCAGCTCAGCCTCATCATCGTCCTGAGCCACAGCCGCCTGTCAGGGGCGGTTCCGGG
CTGGCTGGGCTCGCTGCCGCAGCTCGGCGAGCTGGCACTCTCCAACAACGAGTTCACCGGAGCAATCCCGATGCAG
CTCAGCAACTGCTCCAAGCTCCTGAAGCTGTCCCTCGACAACAACCAGATCAATGGAACAGTGCCGCCTGAACTCG
GCGGCTTGGTGTCCCTCAACGTACTGAATCTCGCACACAACCAGCTCTCAGGTCCGATTCCGACGACGGTCGCAAA
```

GCTGAGCGGCCTCTATGAGCTTAACCTGTCGCAGAATTACCTGTCCGGCCCGATCCCTCCGGATATCGGCAAGTTG
CAAGACTTGCAGAGCCTGCTGGACTTGAGCAGCAACAATCTCAGTGGCCACATCCCTGCATCGCTGCGGCTCACTC
CCCAAGCTGGAAAACCTGAACCTGTCCCACAACGCTCTGGTCGGCGCGGTGCCGTCCCAGCTCGCTGGAATGAGTA
GCTTGGTGCAGCTGGACCTGTCCAGCACCAGCTGGAAGGGAAGCTGGGCACCGAGTTCGGCCGGTGGCCGCAGGCC
GCATTCGCTGACAATACAGGGCTCTGCGGTAGCCCCTTGAGAGGTTGCAGCAGCAGAAACAGCCATTCGGCGCTGC
ACGCGGCGACCATCGCGTTGGTGTCTGCGGTGGTCACGCTGTTGATTGTCCTCCTGATCATTGCGATTGCGCTGAT
GGTGGTGCGCCGCAGGGCCCGGGGTTCAGGCGAGGTGAACTGCACGGCGTTCTTGTCGTCGAGCTCGGGCAGCGCA
AACCGGCAGCTCGTCGTCAAGGGCTCGGCGCGGCGGGAGTTCCGGTGGGAGGCGATCATGGAGGCCACGGCGAACC
TGAGCGACCAGTTCGCCATCGGGTCCGGCGGATCAGGCACGGTGTACAGGGCGGAGCTGTCCACTGGCGAGACGGT
TGCCGTGAAGAGGATAGCGCACATGGACAGCGACATGCTGCTCCACGACAAGAGCTTCGCGCGGGAGGTCAAGATC
CTGGGCCGCGTCCGTCACCGGCACCTGGTCAAGCTGCTCGGCTTCGTCACGTCCCGCGAGTGCGGCGGCGGCGGCG
GCATGCTCGTGTACGAGCACATGGAGAACGGCAGCCTCTACGACTGGCTGCACGGCGGCAGCGATGGCCGGAAGAA
GCGGACGCTCAGCTGGGAAGCGCGGCTCATGGTTGCCGCCGGGCTGGCGCAGGGCGTGGAGTATCTCCACCACGAC
TGTGTGCCCCGCATCGTGCACCGGGACATCAAGTCCAGCAATGTGCTCCTCGACGGCGACATGGAGGCGCACCTCG
GCGACTTCGGCCTCGCCAAGGCCGTCGCCGAGAACCGGCAGGCCGCCTTCGATAAAGACTGCACCGAGTCAGCTTC
CTTCTTCGCCGGATCATACGGGTACATCGCTCCAGGTAATTTCGACGGCAATCTGAAATGCTATAGAAACGCAGTA
GCTCAGGCGACGCGGCCAGTTACTGACAGTGGACGTGCCACATTATCTCTGCAGATGTGCTTACTCCCTGGAGGCG
ACGGAGAGAAGCGACGTCTACAGCATGGGCATCGTGGTTATGGAGCTCGTCACCGGGCTCTTGCCGACCGACAAAC
CTTCGGCGGCGAACATGACATGGTGAGGTGGGTGCAGTCGAGGATGGACGCGCCGTTGCCAGCAAGGGAGCAGGTG
TTCGATCCTGCTCTGAAGCCGCTGGCGCCGCGTGAGGAGTCGTCGATGACGGAGGTGCTG
GAAGTGGCGCTCCGGTGCACAAGGGCTGCTTCACGTTTCACTCGATTACTGACGCGTCTGCGAGATGAATTAGGCT
TTGTTCGTTTGTGTCGGAGTCGAGTGTAATCCAACACAAACAAGACCTTAGGATTTGAAAAGTGGGCTGGATTGTC
TTAGCAGCCGCTCAAAATTGGATCGCTTAATTTCAAATTTTGAGCTCCCTTGTACAATCATCATTTGGAAATGTCC
AGCAG

SEQ ID NO: 16

DM4 cDNA CLONE

GGCACGAGGGGAACTAGACATGTCTGGGAACAAGATTCTGTCGGGCCGAGTACCGGAGTTCCTGGGGGGCTTCCGA
GCATTGCGGCGACTCGGACTTGCCGGGAACAACTTCACCGAGGAAATCCCGGATGAGCTGAGCCTCCTGTGCGGAA
CGCTGGTTCAGCTAGACTTGTCAAGCAACCAATTGGTTGGCGGTTTGCCAGCGAGCTTCTCAGGATGCAGATCGCT
TGAGGTGCTCGATCTGGGTAGCAACCAGCTGTCAGGTGACTTTGTGATCACTGTGATCAGCAAGATCTCTTCTCTG
CGTGTGCTGAGGCTGCCGTTCAACAATATCACAGGCACAAATCCTCTGCCCACGCTAGCAGCTGGCTGCCCTTTGC
TTGAAGTCATTGATCTCGGGTCTAACATGCTGGAANGAGAGATCATGCCCGAGCTGTGTTCATCTTTGCCATCACT
CAGAAAGCTGCTCCTACCCAACAACTACATCAATGGAACCGTGCCGCCCTCACTCGGCAACTGCAGTAATCTGGAG
TCACTGGACCTCAGCTTCAACCTCATGGTTGGTCCGATCACCCTGAAGTACTGTTGCTTCCTAAACTTGTTGAATT
GGTCATGTGGGCAAACANTCTCTCCGGTGAAATACCAAACACGCNATGCTCCACAGCACAACACTGAAAANCCGTC
NTAAACTACAACAACATAACCGAATGATCCCGTTCNCNTCNCCAGTNGCTNAATCCATATGGTGTCCTTCCGGCAA
CACATAACGGGAATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNGGATCCTNTTTCNTNACAGNGGGGATTTTTATATGGTGTAATTGCGGC
AACGGATGACGGGGTNTCCCGCGGTTTGGGACCTTCCAGAGTTTGCCATTTACAGTGCACAGGAATTCATTTTTTG
TTCCTGTGCCAGCAGAGNTTGTTCGCTGCAGCACCTTATNTGGTTGATTTTCAACAGCAACAATTTTTCCGGTGCG
ATACCGCCGCAGTTAGCAGCAACAGGCAGGGNTCATCACTGGAGGCATGGTTTTCTGGGAAGCAGTTCGCGTTCCT
GAGGAATGAGGCTGGAACATCTGCCCAGGTGCTGGAGTGCTGTTCGAGTTCTTTGACATCCGTCCGGAGCGGCTGG
CCCAGTTCCCTGCTGTGCACTCGTGTGCCTCCACGAGGATATACACTGGGATGACAGTGTACACCTTCAACCAAAG
TGGGAGCATGATATTCCTTGATCTCTCGTACAACAGCCTCACAGGCACAATTCCGGCGAGCCTGGGGAACATGACG
TATCTTGATGTCCTTAACCTGGGGCATAATGACCTTACCGGTGCAATTCCAGATGCGTTCACAGGGTTGAAGGCGA
TTGGTGTCCTTGACCTNTCGCACAATCACCTCACCGGTGTCATCCCTGCTGGACTGGGGTGTTTAAATTTCCTAGC
TGACTTCGACGTCTCCAACAACAACCTCACTGGTGAGATACCCACGTCAGGGCAGCTCAGTACATTTCCAGCATCC
CGTTTTGAGAACAACTCCGGCATCTGTGGCATCCCCTTGGATCCTTGCACGCACAATGCCAGCACTGGAGGTGTTC
CTCAAAACCCCTCTAACGTGCGGAGGAAGTTTCTCGAAGAGTTCGTGCTCCTTGCAGTGTCGCTCACCGTGCTCAT
GGTGGCCACCTTGGTTGTCACTGCATACAAGCTCAGGAGGCCCCGTGGGAGCAAAACTGAAGAGATTCAAACTGCT
GGGTATAGCGACAGCCCCGCATCGTCCACCAGTACAAGCTGGAAGCTTTCTGGTTCCAAAGAGCCACTGAGCATCA
ATCTGGCGATATTTGAGAATCCGTTGAGGAAACTAACATATGCCCACCTGCATGAGGCTACCAATGGCTTCAGCTC
GGAAGCTCTTGTTGGCACAGGAGGATTCGGTGAGGTTTACAAGGCTAGGCTCATGGATGGCAGCGTTGTGGCTGTC
AAGAANCTGATGCATTTCACAGGCCAAGGCGACCGGGAGTTCACTGCAGAGATGGAGACCATTGGCAAGATCAAAC
ATCGCAACCTTGTGCCGTTGCTAGGCTACTGCAAAGTTGGCGACGAACGTCTGCTTGTGTACGAGTACATGAATAA
TGGAAGCCTGGATGTCTTGCTCCATGAAAGGGACAAGACTGATGTGGGTCTTGATTGGGCAACAAGGAAGAAGATT
GCAGTTGGCTCGGCAAGAGGACTGGCCTTCCTCCACCATAGTTGCATCCCACACATCATACACCGGGACATGAAGT

CAAGCAACGTGCTTCTTGACGATAATCTCGATGCCTACGTATCGGATTTCGGAATGGCGCGGCTCGTGAATGCTGT
TGACTCACATCTAACCGTGAGCAAGCTCTTAGGAACACCTGGTTATGTGGCTCCCGAGTACTTCCAGTCGGTTATT
TGCACAACTAAGGGCGACGTCTACAGCTATGGCGTTGTTCTTCTGGAGCTTCTCTCAGGGAAAAAACCAATCAATC
CGACTGAATTCGGCGACAATAATCTCATCGACTGGGCCAAGCAGATGGTTAAGGAGGACCGGTGCAGCGAGATATT
TGATCCTATATTGACCGACACAAAATCCTGCGAGTCGGAGCTGTACCAGTATCTGGCGATTGCTTGCCAGTGCTTG
GACGATCAACCTAGTCGCAGACCTACGATGATCCAGGTCATGGCAATGTTCAGTGAGTTTCAGATTGACTCTGGCA
GCTTCTTCTTGGACGGCTTCTCGCTCGATTCAGATAGAGGAATCATCTGAAAAAAAATGTGTAAATGTTATTGATC
CCTGCAGATTATATGATTCACTGGATTTAGGTATTAGCTTAGCCATGTTTAACTCATGTTAACAGGATACAAACAG
ATGTAAATTTGTTTCGGTTGCCGTACATAGTACACAACAGCTTCAACACAGATACCATATAGAGTTGTTTCCAAAA
AAAAA

SEQ ID NO: 17

TRK1

ATCGGGCAGGTCTTCAAAATACTTGTTACATCTTCTTCTTACCTTTGATA
TTTTCCAAAGTATTTGTAACTTCAAATCACTAGTTATCTAAATGGCTACT
TCTAACACAAGTCTCTTGTTTTTCGCGTATTTCCTCCTTGTGTTCCTTAT
TACTCCATCTCAATCGCGTAACCTGTCTCTGAGACGACAGGCTAAAACTC
TAGTTTCATTGAAATATGCATTTGTACAATCATCTGTTCCTAGTACTCTG
TCCAATTGGAACATGTCGAATTATATGTCTATATGTTCTTGGACAGGTAT
AACGTGTGATGATACCAAATCAGTAACTTCCATTGATATATCCAATCTAA
ACATTTCTGGCTCTTTATCACCTGATATTCATGAGCTCACTAGACTTCGC
GTCCTGAATATTTCTAACAATTTGTTTAGTGGAAACTTAAGCTGGGAGTA
TCGCGAGTTTAATGTACTTCAAGTGTTGGATGCTTATAACAACAATTTCT
CTGGTCCACTCCCTTTGGGAGTTACTCAACTTGTGCAGCTCAAGTACTTG
AATTTCGGGGGTAACTACTTTTCAGGGAAGATTCCTTTGAGTTATGGTAG
TTTTAATCAGCTTGAGTTCCTGTCTCTTGCTGGGAATGACTTGCACGGTC
CTATACCGAGGGAGCTGGGGAACGTTACGAGCCTCAGGTGGTTACAGTTG
GGTTATTATAATCAATTTGATGAGGGGATTCCACCAGAGTTGGGGAAACT
TGTTAATTTGGTTCATCTAGATCTTTCAAGCTGTAACTTAACGGGTTCGA
TTCCACCAGAATTGGGCAATCTTAATATGTTGGACACTCTTTTCTTGCAA
AAGAATCAACTTACTGGTGTATTTCCTCCTCAGCTAGGGAATTTGACAAG
GTTAAAATCTCTTGATATCTCGGTCAATGAACTCACAGGAGAGATCCCGG
TTGACTTGTCAGGACTCAAGGAGCTCATATTGTTGAACCTCTTTATCAAC
AATTTGCACGGTGAGATTCCAGGATGTATCGCGGAGCTGCCAAAGTTGGA
AATGTTGAATCTTTGGAGGAATAATTTCACTGGCTCGATTCCTTCTAAGC
TTGGGATGAACGGTAAACTAATTGAAATTGATCTGTCTAGTAATAGACTC
ACTGGCTTGATACCAAAATCTCTATGCTTTGGGAGGAATTTGAAAATCTT
GATTCTTCTTGATAATTTTCTGTTTGGACCTTTACCTGATGATTTTGGGC
AGTGTCGAACGTTGTCCAGAGTCAGAATGGGACAGAATTACTTGAGTGGA
TCAATACCAACAGGGTTTCTTTATTTGCCTGAGTTGTCACTGGTGGAACT
GCAGAACAACTACATCAGTGGACAACTCTGGAACGAGAAAAGCTCAGCGT
CTTCTAAACTTGAAGGGCTGAACCTGTCGAACAATCGCTTGTCTGGTGCA
CTTCCTAGTGCTATTGGAAACTATTCAGGGCTGAAGAATCTTGTGTTAAC
TGGAAATGGTTTCTCAGGTGATATCCCTTCTGATATTGGCAGACTAAAGA
GCATCTTAAAGCTGGACCTGAGTAGAAACAACTTCTCTGGCACAATCCCT
CCTCAGATTGGTAACTGTCTTTCCTTAACTTACTTGGATTTGAGCCAAAA
TCAACTTTCTGGTCCTATCCCAGTTCAAATTGCTCAAATTCACATCTTAA
ATTACATCAATATTTCCTGGAATCACTTCAACGAGAGCCTTCCCGCGGAG
ATTGGCTTGATGAAGAGTTTAACTTCAGCAGATTTTTCCCACAATAACTT
ATCTGGATCAATACCTGAAACAGGCCAATATTTATATTTCAACTCAACTT
CCTTCACCGGCAACCCTTATCTCTCTGGATCCGACTCGACTCCTAGCAAC
ATTACATCCAACTCACCGTCAGAACTTGGAGACGGAAGTGACAGCAGAAC
TAAGGTTCCTACAATATACAAGTTCATATTTGCATTTGGGCTCTTATTCT
GCTCCCTCATTTTCGTTGTCTTAGCAATAATCAAGACAAGAAAGGGGAGT
AAGAATTCAAATTTGTGGAAGCTGACAGCATTTCAGAAGCTTGAGTTCGG
AAGTGAAGACGTCTTGCAGTGCTTGATAGACAACAACGTCATAGGGAGAG
GTGGAGCAGGGATAGTGTATAAGGGAACTATGCCAAATGGTGATCATGTC
GCGGTGAAGAAATTGGGAATAAGCAAAGGCTCACATGATAACGGCCTATC

TGCTGAACTTAACACATTAGGGAAGATCAGGCATAGGTACATTGTGAGAC
TGCTCGCGTTTTGTTCAAACAAGGAAGTCAACTTGCTAGTTTATGAGTAC
ATGCTAAATGGAAGCTTAGGTGAAGTGCTTCATGGGAAGAACGGCGGGCA
ACTCCAATGGGAAACTAGGCTAAAAATAGCCATAGAAGCTGCCAAGGGCC
TTTCTTATTTGCACCACGATTGCTCCCCTATGATAATCCACCGCGATGTC
AAGTCCAACAATATATTGTTGAACTCTGAACTTGAAGCTCATGTTGCAGA
TTTTGGATTAGCCAAGTACTTTCGTAACAATGGTACCTCTGAGTGCATGT
CTGCAATTGCAGGATCTTATGGCTACATTGCTCCAGAATATGCATACACG
CTGAAAATTGATGAGAAAAGCGATGTGTATAGCTTTGGAGTGGTGTTGTT
GGAGCTTATAACAGGACGAAGGCCAGTAGGAAATTTTGGAGAAGAAGGAA
TGGACATTGTACAATGGGCGAAAACGGAGACAAAATGGAGCAAAGAAGGG
GTGGTGAAAATCTTGGATGAGAGGCTAAAAAATGTTGCAATTGTTGAAGC
TATGCAAGTATTTTTTGTAGCAATGCTTTGTGTTGAAGAGTACAGCATTG
AGAGGCCTACAATGAGGGAAGTAGTCCAAATGCTTTCTCAAGCTAAACAA
CCAAATACTTTCCAAATCCAATAATCTAATTGTGGCTCTACTTATTGTAT
GCTTGGGAACACCCCTTTTGTTAGCTTTGCAAAAGTGAAATCACAAATTA
ATCTAAGTGAAGTAGTTGCAAAATTAATTTGCAATTATGTTAGATCTTAG
GGTATGATATCTAACTATATCCTCTCAACTTGGAATAGTGTATTGGATGT
GTAGAACTAGTATTAGTATCCGCGTGATGTGTGGCGAATATCAAAAGAAA
GTCGACRCAKRMMRCTAATYCCKCYGWTWCWRAMGKMSWCCMGGMSKCGA
WYKCCRCCRATACTGACGGACTCCAGGAGTCGTCGCCACCAAT

SEQ ID NO:18
TRK1
SGRSSKYLLHLLLTFDIFQSICNFKSLVI.
MATSNTSLLFFAYFLLVFLITPSQSRNLSLRRQAKTLVSLKYAFVQSSVPSTLSNWNMSNYMSICSWTGITCDDT
KSVTSIDISNLNISGSLPPD
IHELTRLRVLNISN NLFS GNLPWE
YREFNVLQVLDAYN NNFS GPLPLG
VTQLVQLKYLNFGG NYFS GKIPLS
YGSFNQLEFLSLAG NDLH GPIPRE
LGNVTSLRWLQLGYYNQFDEGIPPE
LGKLVNLVHLDLSSCNLT GSIPPE
LGNLNMLDTLFLQK NQLT GVFPPQ
LGNLTRLKSLDISV NELT GEIPVD
LSGLKELILLNLFI NNLH GEIPGC
IAELPKLEMLNLWR NNFT GSIPSK
LGMNGKLIEIDLSS NRLT GLIPKS
LCFGRNLKILILLD NFLF GPLPDD
FGQCRTLSRVRMGQ NYLS GSIPTG
FLYLPELSLVELQN NYIS GQLWNEK
SSASSKLEGLNLSN NRLS GALPSA
IGNYSGLKNLVLTG NGFS GDIPSD
IGRLKSILKLDLSR NNFS GTIPPQ
IGNCLSLTYLDLSQ NQLS GPIPVQ
IAQIHILNYINISW NHFN GSLPAE
IGLMKSLTSADFSH NNLS GSIPET

GQYLYFNSTSFTGNPYLSGSDSTPSNITSNSPSELGDGSDSRTKVPTIYK
FIFAFGLLFCSLIFVVLAII
KTRKGSKNSNLWKLTAFQKLE
FGSEDVLQCLKDNNVIGRGGAGIVYKGTMPNGDHVAVKKLGISKGSHDNGLS
AELNTLGKIRHRYIVRLLAFCSNKEVNLLVYEYMLNGSLGEVLHGKNGGQLQ
WETRLKIAIEAAKGLSYLHHDCSPMIIHRDVKSNNILLNSELEAHVADFGLA
KYFRNNGTSECMSAIAGSYGYIAPEYAYTLKIDEKSDVYSFGVVLLELITGR
RPVGNFGEEGMDIVQWAKTETKWSKEGVVKILDERLKNVAIVEAMQVFFVAM
LCVEEYSIERPTMREVVQMLSQAKQPNTFQIQ

SNCGSTYCMLGNTPFVSFAKVKSQINLSEVVAKLICNYVRS.GMISNYILS
TWNSVLDV.N.Y.YPRDVWRISKESR???LIP??????R?R?PPILTDSRSRRHQ

SEQ ID NO: 19

TRK2

TGACTCTCTCTGTCTCTCTCTGTTCGCAGCCCCAAAAAGTAGGGTTAGGG
CTAGGGTTTTTGAGTTTCAAAACCCCATTTCTGGTTCCTATAATCTTCAC
ATACAAGGGGAGTTTGTCTCTGTTGCATTCTTTGAAGACCCTTTTGGGGT
TTTACTAATGGGTCGTTGTTGTTTTGTCATCAAATGGTACTATCATGACA
TACCCTTGAAAGTTTTTCTCATCCTTTGTGTTTTCTTCTTAGTTCATGGC
TATGCACTTTCCTCGGATTCGGATAAATCAGCGCTCTTGGAGTTAAAGGC
CTCATTTTCAGATTCCTCTGGAGTGATTTCTAGCTGGAGCTCCAGAAATA
ATGATCACTGTTCATGGTTTGGTGTCTCCTGTGATTCCGATTCACGTGTT
GTGGCTTTGAACATCACTGGAGGTAATTTGGGTTCTTTATCTTGTGCTAA
AATTGCTCAATTTCCTTTGTATGGCTTTGGAATTACAAGGGTTTGTGCTA
ATAATAGTGTCAAGCTTGTTGGTAAAGTACCTCTCGCAATATCAAAATTA
ACTGAACTAAGGGTTTTATCCTTGCCTTTTAATGAATTGCGTGGTGATAT
TCCATTGGGAATTTGGGATATGGACAAACTTGAAGTTTTGGATCTGCAAG
GGAATTTAATTACTGGGTCTTTGCCATTGGAGTTTAAGGGGTTGAGGAAA
TTGAGGGTTTTAAACTTGGGTTTTAATCAGATTGTGGGTGCCATACCGAA
TTCCTTGTCAAATTGCCTTGCTCTACAAATCTTTAATCTTGCTGGAAATA
GGGTAAATGGGACCATTCCAGCATTCATTGGTGGATTTGAAGATCTGAGG
GGAATCTACCTGTCTTTTAATGAGCTTAGCGGGTCTATTCCTGGTGAAAT
TGGGCGTTCTTGTGAGAAGCTTCAAAGTCTAGAGATGGCAGGTAATATCT
TAGGTGGTGTTATTCCAAAAAGTTTAGGGAACTGCACACGGTTGCAGTCA
CTTGTCTTATATTCAAATTTGTTGGAAGAGGCTATTCCAGCTGAATTTGG
TCAACTAACTGAGCTCGAGATTCTTGATTTGTCTAGGAACAGCCTAAGTG
GTCGACTACCATCTGAGCTGGGAAACTGCTCGAAACTATCCATTCTTGTA
CTGTCAAGTTTGTGGGATCCCCTTCCAAATGTGTCTGATTCAGCTCATAC
TACTGATGAGTTTAACTTTTTTGAAGGCACAATCCCATCAGAGATCACCA
GGCTTCCTAGTTTGAGAATGATATGGGCTCCCAGGTCAACTCTTTCAGGA
AAATTTCCTGGCAGTTGGGGTGCTTGTGACAATTTGGAGATCGTGAACTT
GGCTCAAAATTATTATACTGGAGTGATTCCTGAGGAATTGGGTAGCTGCC
AGAAGTTGCATTTTCTTGACTTGAGCTCAAATAGGCTGACTGGACAGCTT
GTTGAGAAACTGCCAGTCCCTTGCATGTTTGTGTTCGATGTGAGTGGGAA
TTATCTCTCTGGTTCAATTCCCAGGTTTTCCAATTACAGTTGTGCTCATG
TTGTTTCCAGCGGTGGAGAGCCATTTGGGCCCTATGATACATCATCTGCA
TATCTAGCACATTTCACCAGTAGAAGTGTTCTAGACACTACATTATTTGC
AGGTGATGGTAACCATGCAGTATTTCATAATTTCGGTGTTAACAACTTCA
CGGGAAATTTACCGCCTTCCATGCTAATTGCACCTGAAATGTTAGGCAAA
CAAATTGTATACGCATTTCTTGCTGGTAGTAACAGGTTTACTGGACCTTT
TGCTGGTAACTTGTTCGAGAAATGTCATGAATTGAATGGAATGATTGTTA
ATGTAAGCAATAATGCGTTGTCAGGTCAAATCCCAGAGGATATTGGTGCA
ATTTGTGGGTCTCTTAGGCTGTTGGATGGATCCAAAAATCAGATTGTTGG
GACAGTCCCTCCGAGTTTAGGGAGTCTGGTTTCATTAGTTGCTCTCAATT
TAAGTTGGAACCACCTGCGAGGTCAGATTCCTAGCAGACTTGGCCAGATA
AAGGATCTCAGTTACCTCTCTTTGGCTGGCAATAATCTGGTTGGCCCAAT
CCCCTCAAGTTTTGGCCAATTGCACTCTTTAGAAACGCTTGAACTTTCTT
CGAATTCTTTGTCTGGTGAAATTCCAAATAATCTGGTAAATTTGAGGAAT
TTGACTTCCCTTCTTCTGAACAACAACAATTTATCAGGGAAAATACCTTC
AGGCTTGGCCAATGTGACCACACTGGCAGCATTTAACGTTTCTTTCAATA
ATCTGTCTGGGCCACTGCCTCTTAACAAAGATTTGATGAAGTGTAATAGT
GTTCAGGGAAACCCCTTTCTGCAATCGTGCCATGTATTTTCTCTATCAAC
ACCTTCTACAGATCAGCAGGGAAGAATAGGGGACTCACAAGATTCTGCTG
CGTCTCCTTCAGGTTCAACCCAGAAAGGAGGGAGCAGCGGTTTCAACTCC
ATAGAGATTGCATCCATAACATCTGCGGCAGCTATTGTGTCAGTTCTTCT
TGCTCTGATAGTCCTGTTCTTTTACACCAGAAAATGGAATCCAAGATCTA
GAGTTGCTGGATCTACCAGGAAAGAAGTCACAGTGTTTACAGAAGTTCCG

GTTCCTTTAACATTTGAAAATGTAGTGCGGGCCACAGGGAGCTTCAATGC
AAGCAATTGCATAGGCAGTGGAGGTTTTGGAGCAACATACAAAGCGGAGA
TTGCACCAGGGTTCCTAGTGGCAGTAAAGCGACTTGCTGTAGGACGTTTT
CAGGGGATTCAACAGTTTGATGCAGAAATCAGAACTCTGGGGAGGCTTCG
ACATCCAAACCTCGTAACTCTGATAGGATATCATAATAGTGAAACAGAAA
TGTTTCTGATCTATAACTATTTGCCAGGTGGTAATTTGGAAAAGTTTATT
CAGGAGAGGTCTACAAGGGCTGTGGACTGGAGGGTTCTTCACAAGATTGC
TTTGGATGTAGCCCGTGCACTTGCTTACCTGCATGATCAGTGTGTACCAC
GTGTGCTTCATCGTGATGTGAAGCCGAGCAACATTTTATTGGATGAGGAG
TATAATGCATATTTATCTGATTTTGGTTTGGCTAGATTACTGGGAACTTC
AGAGACCCATGCAACTACTGGTGTGGCGGGAACTTTTGGATATGTTGCTC
CTGAATATGCCATGACTTGCCGCGTCTCGGACAAGGCTGATGTCTACAGT
TATGGGGTTGTGTTGCTTGAGTTAATATCAGATAAGAAAGCACTAGATCC
GTCTTTCTCTTCTTATGGAAATGGATTCAATATTGTAGCTTGGGCATGCA
TGCTTTTACGCAGGGCCGTGCTAAGGAGTTCTTTACGGCTGGTCTATGGG
ATTCAGGTCCACATGATGATTTGGATGAGGTCCTACACTTGGCAGTGGTC
TGCACGGTTGACTCTCTTTCTACTAGACCAACAATGAAGCAAGTAGTAAG
ACGGTTGAAGCAACTTCAACCCCCGTCGTGTTAGCTGCGGCATGTGTTTT
GGATAGGATATGGTTTAGCCCAATTGTAATNTTAAAACTTGCCCTTGATA
GTAAGGTGTATTTGGGTGTCTCGTATTAGGTTCAGATTTGTATTTGTAGC
CTGCTTGTGAATTGTAGTATATAGCCAGCCCCC:ATTTTTCC:ATGTCAT
GTCCC:TAATTAGGGGGTGTGCAGATTCTTCT:GCAGAAGAGTGCAGATA
CTTGTCTTCAACATGTACC:ACATTTTTTTTTGTTTGTTAAATAAGAGCA
AAAAATAGGAACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA:::GC

SEQ ID NO: 20
TRK2

DSLCLSLFAAPKSRVRARVFEFQNPISGSYNLHIQGEFVSVAFFEDPFGVLLMGR
CCFVIKWYYHDIPLKVFLILCVFFLVHGYALSSDSDKSALLELKASFSDSSGVIS
SWSSRNNDHCSWFGVSCDSDS
RVVALNITGGNLGSLSCAKIAQFPLYGFG
ITRVCANNSV
KLVGKVPLA
ISKLTELRVLSLPF NELRGDIPLG
IWDMDKLEVLDLQG NLITGSLPLE
FKGLRKLRVLNLGF NQIVGAIPNS
LSNCLALQIFNLAG NRVNGTIPAF
IGGFEDLRGIYLSF NELSGSIPGE
IGRSCEKLQSLEMAGNILGGVIPKS
LGNCTRLQSLVLYS NLLEEAIPAE
FGQLTELEILDLSR NSLSGRLPSE
LGNCSKLSILVLSSLWDP
LPNVSDSAHTTDEF NFFEGTIPSE
ITRLPSFENDMAPR STLSGKFPGS
WGACDNLEIVNLAQ NYYTGVIPEE
LGSCQKLHFLDLSS NRLTGQLVEK
LPVPCMFVFDVSG NYLSGSIPRF
SNYSCAHVVSSGGEPFGPYDTSSAYLAHFTSRSVLDTTLF
AGDGNHAVFHNFGV NNFTGNLPPS
MLIAPEMLGKQIVYAFLAGSNRFTGPFAGNL
FEKCHELNGMIVNVSNNALSGQIPED
IGAICGSLRLLDGSKNQIVGTVPPS
LGSLVSLVALNLSW NHLRGQIPSR
LGQIKDLSYLSLAG NNLVGPIPSS
FGQLHSLETLELSS NSLSGEIPNN
LVNLRNLTSLLLNN NNLSGKIPSG
LANVTTLAAFNVSF NNLSGPLPLN

KDLMKCNSVQGNPFLQSCHVFSLSTPSTDQQGRIGDSQDSAASPSGSTQKGGSSGFNSIEIASITSAAAIVSVLLA
LIVLFFYT
RKWNPRSRVAGSTRKEVTVFTEVPVPLTFE
NVVRATGSFNASNCIGSGGFGATYKAEIAPGFLVAVKRLAVGRFQGIQQFDAEIRTLG
RLRHPNLVTLIGYHNSETEMFLIYNYLPGGNLEKFIQERSTRAVDWRVLHKIALDVAR
ALAYLHDQCVPRVLHRDVKPSNILLDEEYNAYLSDFGLARLLGT
SETHATTGVAGTFGYVAPEYAMTCRVSDKADVYSYGVVLLELIS
DKKALDPSFSSYGNGFNIVAWACMLLRRAVLRSSLRLVYGIQVH
MMIWMRSYTWQWSARLTLFLLDQQ
.SK..DG.SNFNPRRVSCGMCFG.DMV.PNCN?KTCP...GVFGCLVLGSDLYL.PACE
L.YIASP?FF?CHV?.LGGVQIL?AEECRYLSSTCT?IFFCLLNKSKK.EPKKKKKKKKK??

SEQ ID NO: 21
TRK3
TCAAGCAACAATTTGTCAGGCCAGATTCCCAAGTCCTTGAGGAATCTTGAACATCTCATGTATTTCAATGTCTCGT
TCAATGGGCTCATGGGTGAAATTCCAGATGGAGGGCCATTCGTAAATTTTACAGCTGAATCATTCATGGGTAACCC
TGCATTATGTGGATCATCACGCTTCCGTGTGATGCAATGCAGAGTCACTAGTCTTGAAAGAAAAGGAAAGAGTAGA
GTCTTAACTTCTGTTCTTGCATCAGCTTCCTCAGGAGTTGTAGTCACGACCATTTTCATCATTTGGTTTCTGAAAT
GCCGAAAAAGGAGTACGGAACTTCCTCTAGTTGATACATTTGGTCAGGTACATAAGAGGATTTCGTACTATGATAT
TCCTCAAGGGACAAACAGCTTTGATGAAGCAAACTTGATTGGAAGGGGGAGCCTTGGTTTGGTCTACAAAGGAAAG
CTTGAAAATCCTAAGCG

SEQ ID NO: 22
TRK3

SSNNLSGQIPKSLRNLEHLMYFNVSFNGLMGEIPDGGPFVNFTAESFMGNPALCGSSRFRVMQCRVTSLERKGKSR
VLTSVLASASSGVVVTTIFIIWFLKCRKRSTELPLVDTFGQVHKRISYYDIPQGTNSFDEANLIGRGSLLGLVYKG
KLENPKR

SEQ ID NO:23
TRK4
AGCTTGGATCAATTTAACAATGATAAACTGCAATTTAGCTGGTCCTTTGCCTGAATTTCTTGGAACTATGTCTTCT
TTAGAGGTTTTGTTGTTGTCTACAAATAGGCTTTCAGGGCCTATTCCAGGTACTTTCAAGGATGCAGTGCTGAAGA
TGCTTTGGTTGAACGATCAGTCGGGTGATGGTATGAGTGGTTCAATAGATGTTGTTGCAACTATGGTATCACTTAC
ACATCTTTGGCTTCATGGGAATCAATTTTCAGGTAAAATCCCAGTAGAGATTGGTAATCTAACAAATCTGAAGGAT
CTCAGTGTGAATACAAATAACCTTGTTGGATTAATCCCTGAAAGTTTAGCTAATATGCCATTAGACAATCTTGATT
TGAATAATAATCATTTTATGGGACCAGTTCCTAAGTTCAAGGCTACTAATGTTAGTTTTATGTCCAACTCTTTTTG
TCAAACCAAACAAGGAGCAGTATGTGCCCCTGAGGTTATGGCACTTTTAGAGTTTCTTGATGGGGTGAATTATCCT
TCTAGGCTTGTTGAATCATGGTCTGGAAACAACCCTTGTGACGGACGTTGGTGGGGAATAAGCTGTGACGATAACC
AAAAAGTTAGTGTTATAAACTTGCCCAAGTCTAATCTTTCCGGGACCTTGAGTCCTTCAATCGCGAACCTTGAAAC
CGTTACTCACATTTATCTTGAATCAAATAATCTTTCTGGTTTTGTTCCATCTAGTTGGACTAGTTTGAAATCTCTG
TCTATTCTTGATTTGAGTAATAACAATATTTCCCCACCTTTGCCTAAATTTACCACCCCTTTGAAACTTGTTCTAA
ATGGAAATCCAAAGCTGACTTCTAATCCTCCTGGAGCAAATCCTTCACCAAACAACAGCACAACTCCTGCAGATTC
ACCCACGTCGTCTGTACCATCTTCACGACCCAACAGTTCAAGCTCTGTGATCTTTAAACCCAGTGAACAGTCACCC
GAGAAAAAGGACTCAAAGTCAAAGATAGCTATAGTTGTGGTTCCTATTGCTGGTTTTCTACTTTTGGTTTGTCTTG
CTATTCCACTGTACATTTATGTCTGTAAGAAGAGTAAAGATAAGCATCAAGCTCCAACTGCTCTTGTGGTTCATCC
TAGAGATCCGTCTGATTCGGATAATGTAGTCAAGATTGCGATTGCCAATCAGACTAATGGAAGTCTTTCCACAGTA
AATGCAAGTGGTTCTGCTAGCATACACAGTGGTGAATCCCATTTGATCGAAGCTGGGAATTTGCTCATATCGGTTC
AAGTACTTCGGAATGTGACTAAGAATTTTTCTCCGGAAAATGAACTTGGACGTGGTGGTTTTGGTGTGGTTTATAA
GGGAGAATTAGATGATGGGACACGAATCGCTGTCAAAAGAATGGAGGCTGGTATTGTTAGCAACAAAGCT

SEQ ID NO:24
TRK4 aa
AWINLTMINCNLAGPLPEFLGTMSSLEVLLLSTNRLSGPIPGTFKDAVLKMLWLNDQSGDGMSGSIDVVATMVSLT
HLWLHGNQFSGKIPVEIGNLTNLKDLSVNTNNLVGLIPESLANMPLDNLDLNNNHFMGPVPKFKATNVSFMSNSFC
QTKQGAVCAPEVMALLEFLDGVNYPSRLVESWSGNNPCDGRWWGISCDDNQKVSVINLPKSNLSGTLSPSIANLET
VTHIYLESNNLSGFVPSSWTSLKSLSILDLSNNNISPPLPKFTTPLKLVLNGNPKLTSNPPGANPSPNNSTTPADS
PTSSVPSSRPNSSSSVIFKPSEQSPEKKDSKSKIAIVVVPIAGFLLLVCLAIPLYIYVCKKSKDKHQAPTALVVHP

RDPSDSDNVVKIAIANQTNGSLSTVNASGSASIHSGESHLIEAGNLLISVQVLRNVTKNFSPENELGRGGFGVVYK
GELDDGTRIAVKRMEAGIVSNKA

SEQ ID NO:25

TRK5 3'

TGAGTCAGCATTTGTGTTGAATGGCAAGAATTAGGCTAACAAACCTCTTAACTTGGAAGGCAAAGGGTAACGATAG
CATTGGATGTGGCAAGAGGGGTTGAATACCTTTCATAGCTTGGCACAACAAAGTTTCATTCATAGAGATTTAAAAC
CGTCAAACATCCTTCTCGGGAGACGACATGAGAGCCAAGGTTGCAGATTTTGGTTTGGTTAAGAAATGCCCCTGAT
GGAAAGTATTCTGTGGAGACACGTTTGGCTGGAACTTTTGGCTATCTTGCACCTGAATACGCGGGTAAGCATTTCC
ACTCCTTGTATCATTACATCTTTTTTAACCGAGAACTGATTTGTATGCCCCTAGCTGTAAACTTTAGATCTTAGTA
AATTTTAAAATTATTCGATCATCTGCATTGCTAGGCATTCTGGTCTTCATTACATTTAGAGACATTTGGCACATAG
TGTTAATACTGAGGCAATCTCCATGCGTAAATGCTGCTTTGTCATTTTTATCGAACTCATGAACTATATGCTACTA
TTATAAAAGATAAGTTTCCATTTCGTGTGGATCAGCATCTGAATTATTATCCTATTACATCTCCTCTTTGTTACAG
CTACTGGACGAGTCACAACCAAAGTAGACGTTTACCTTTGGCGTTGTTTTGATGGAGATCATTACTGGTAGAAAAG
CT

SEQ ID NO:26

TRK5 5'

CTTTCTCCAACTCCTTCTGGTTGGTAAGTTCTAAGGCTTTCTGTTCTTGGACTAATGTATTTGTGACAAGTCTTCA
TCTACAGTAACTTCAATTAATCTTGATTCTCAATCTGTTTCTGGTTCTTTACCTTCTGANATTAGTCAACTTTCTA
ATCTTAAAACCCTTTCACTTCAAAAAAACAAACTTTCTGGCCCTTTACCTTCTTTTGCCAACATGTCAAAATTAGC
TGATCTTTTCTTGGACAATAACCAATTCACTTCTGTTCCTCAAGATTTCCTTTTGGGGGTTCCTAGTTTAGTAACT
TTAAGCATTANTGAAAATGCGGGACTCTCTCCTTGGCAAATACCTATGTATTTAACTGANAGTACCAAATTTGGGA
TCTCTATATGCTAGTAATGCAAGTATT

SEQ ID NO:27

TRK5 5' aa

LSPTPSGWSTSKPFCSWTNVICDKSSSTVTSINLDSQSVSGSLPSDISQLSNLKTLSLQKNKLSGPLPSFANMSKL
ADLFLDNNQFTSVPQDFLLGVPSLVTLSISENAGLSPWQIPMYLTESTKFGISIC..CKY

SEQ ID NO:28

TRK6 3'

GCTTATGACAGTCATTGATCCAACTCTAGACGTGAAAGATGAGATAACCGAGAGCATTTCCACCTTAGCTGAACTT
GCTGGTCATTGCACTGCAAGAGAACCTGGTCAACGGCCAGATATGGGCCACGCTGTAAACGTGCTGTCCCCACTTG
TTGAGAAATGGAAGCCTCTTGAGGATGATCCAGCAGGACTATTGTGGTATCGACTACAGTCTTCCCCTCAATCAAA
TGGTCAAGGGTTGGCAAGAATCGGAAGGAAAAGACTTAAGTTACGTGGATCTCGAGGACAGTAAGGGCAGTATCCC
AGCAAGACCAACTGGATTTGCAGATTCATTTACATCAGCTGATGGTAGATAATGGAGGTACTTCTATGTAGTAGAT
GTAGATATCAATTTTCTTTGTATTGTATTGAGATTTTGATCGTATTTTCCACGTGCCTTCGCTCATTTCTCCCCCT
TCAATGTGAATGTATTAGAAATTTAAACTATGTGTAGCCTCAGTTCCTTCTGTAGATATAAAATAGCGGTGAAGGA
GAACTATACATCGATCAGTTGAACTCCTCGCTCAGTCACTATTTTCATCTTCTACTATGGTGAGATTTAAGAGCAT
TTTTTTCACCTTTGCCTATTTTCGTCGTTTAGTTGTGCTGCTTCTGAAGCTCTGTGCTGAATACATACTGAGCGGT
GAAGTAGCCCGGTATATAGAATATGTCGTTTGATTCGAAGAATCTAACCATATAGCTTCTTCCTTCTGCTCAAATA
GGGCAATACTTGGAGAAGATAACCCTTATGTGTCCTCATCATTTGTTGTTTTGAGAAAAAAGTATGATATATGTTG
CCTTAGACTCTTCGAAAAGGTATTGTCAT

SEQ ID NO:29
TRK7 3'

AAGCTTTCGATTCTCCCTATATGTCTCTCATGTTAACTAATATAGCTTGATAAACCTAGTAAATTCTTCATACTTT
GTGTATATAGTGATATCGAAACGTGAGATGCAGCCTTTATCCGCCTTGTTCTTTTTTCTTGGCACTCTACTTGTTC
TGTTACTTCTCATATTTATCTTATGCAAGTTTTTCAAACCAGGAAAACTGACAGCTTTGATTACTAGCACATTTAA
GCCTCTTCCAGGTTTTGTTTTGTTTCGAATTCATGTTACTTTGAGTATCTTATTCGTTATTTTAAAATATGTCTTT
TTTCGGATGATACAGAGATGAGTAATGCTGTTAGTGTACATCTTCATACAATAAGCTACTTTAACTTTCATACATT
GGAGAAAGCTACGAAGAATTTTCACTCGGATAACCTCCTTGGATGTGGTGGATTTGGTCCGGTTTTCCTGGTAAGG
TCAATGTGTCTTAACGAGTTGTAAAATCCGTTTTTATAATTAGATGCAGAGGTTAAGAATTGGCCCTAGTGTTTTT
TCTTAGGGGAAGTTAGGAGATGGACAATTAGTTGCTGTCAAGAAGTTGTCTGTTGACAAATCGCAGCAAGGGGACC
GGGAGTTTCTTGCTGAGGTACGGATGATAACAAGTATACAACACAAGAACCTTGTCCGCCTTCTTGGTTGTTTGTT
CAAAAGGGGGATCAGAGGCCTTCTCGTGTAACGAATAACATGAAGAATAAGGAGGCTTGGGACCAAATACTATATT
GGTATTTAT